

# THE STAR FORMATION EPOCH OF THE MOST MASSIVE EARLY-TYPE GALAXIES<sup>1</sup>

PIETER G. VAN DOKKUM<sup>2,3</sup> AND ROELAND P. VAN DER MAREL<sup>4</sup>

Received 2006 July 19; accepted 2006 September 19

## ABSTRACT

We present new spatially resolved Keck spectroscopy of early-type galaxies in three galaxy clusters at  $z \approx 0.5$ . We focus on the fundamental plane (FP) relation and combine the kinematics with structural parameters determined from *HST* images. The galaxies obey clear FP relations, which are offset from the FP of the nearby Coma Cluster due to passive evolution of the stellar populations. The  $z \approx 0.5$  data are combined with published data for 11 additional clusters at  $0.18 \leq z \leq 1.28$ , to determine the evolution of the mean  $M/L_B$  ratio of cluster galaxies with masses  $M \gtrsim 10^{11} M_\odot$ , as implied by the FP. We find  $d \log (M/L_B)/dz = -0.555 \pm 0.042$ , stronger evolution than was previously inferred from smaller samples. The observed evolution depends on the luminosity-weighted mean age of the stars in the galaxies, the IMF, selection effects due to progenitor bias, and other parameters. Assuming a normal IMF but allowing for various other sources of uncertainty, we find  $z_* = 2.01^{+0.22}_{-0.17}$  for the luminosity-weighted mean star formation epoch. The main uncertainty is the slope of the IMF in the range  $1\text{--}2 M_\odot$ : we find  $z_* = 4.0$  for a top-heavy IMF with slope  $x = 0$ . The  $M/L_B$  ratios of the cluster galaxies are compared to those of field early-type galaxies at  $0.32 \leq z \leq 1.14$ . Assuming that progenitor bias and the IMF do not depend on environment, we find that the present-day age of stars in massive field galaxies is  $4.1\% \pm 2.0\%$  ( $\approx 0.4$  Gyr) less than that of stars in massive cluster galaxies. This relatively small age difference is surprising in the context of expectations from “standard” hierarchical galaxy formation models and provides a constraint on the physical processes that are responsible for halting star formation in the progenitors of today’s most massive galaxies.

*Subject headings:* cosmology: observations — galaxies: evolution — galaxies: formation

## 1. INTRODUCTION

Recent studies of near-infrared selected samples have uncovered a large population of red, massive galaxies at  $z > 2$  (e.g., Labbé et al. 2003; Franx et al. 2003; van Dokkum et al. 2003; Daddi et al. 2004; Yan et al. 2004; Papovich et al. 2006). Red galaxies dominate the high-mass end of the mass function at  $2 < z < 3$  (van Dokkum et al. 2006) and are highly clustered (Daddi et al. 2003; Grazian et al. 2006; Quadri et al. 2007). Most of these objects are too faint in the rest-frame ultraviolet to be selected as Lyman break galaxies (e.g., Förster Schreiber et al. 2004; Reddy et al. 2005). The most straightforward interpretation of these newly found objects is that they are progenitors of today’s massive early-type galaxies.

Interestingly, a large fraction of these galaxies appear to have very high star formation rates, based on modeling of their spectral energy distributions, (stacked) X-ray emission, infrared emission, submillimeter emission, and spectra (e.g., Förster Schreiber et al. 2004; Rubin et al. 2004; Daddi et al. 2004; Labbé et al. 2005; Knudsen et al. 2005; Papovich et al. 2006; Webb et al. 2006). A similar or smaller fraction is best fitted by passively evolving models with little ongoing star formation (e.g., Labbé et al. 2005; Papovich et al. 2006; Kriek et al. 2006b). These results suggest that by  $z \sim 2.5$  we are entering the star formation epoch of massive early-type galaxies.

An important question is whether the properties of early-type galaxies at intermediate redshift are consistent with this interpre-

tation. Furthermore, detailed studies of early-type galaxies in different environments may identify descendants of particular populations of massive galaxies at  $z > 2$ . For example, it may be that the passively evolving galaxies at  $z \sim 2.5$  are progenitors of cluster early-type galaxies whereas the star-forming galaxies are progenitors of early-type galaxies in groups.

One of the most sensitive tools for determining the star formation epoch of early-type galaxies is the redshift evolution of the fundamental plane (FP) relation (Djorgovski & Davis 1987), as it reflects evolution in the  $M/L$  ratios of galaxies and has very low scatter. The  $z > 0$  FP has been studied extensively in the past decade, both in clusters (e.g., van Dokkum & Franx 1996; van Dokkum et al. 1998a; Wuyts et al. 2004; Fritz et al. 2005; Holden et al. 2005; Jørgensen et al. 2006) and in the field (e.g., Treu et al. 1999, 2002, 2005a; van Dokkum et al. 2001; Rusin et al. 2003; van de Ven et al. 2003; van der Wel et al. 2004, 2005; di Serego Alighieri et al. 2005). The results from these papers can be summarized as follows: (1) the  $M/L$  ratios of the most massive cluster and field early-type galaxies evolve slowly and regularly, indicating early formation of their stars; and (2) there is evidence that low-mass galaxies evolve faster than high-mass galaxies, both in the field and in clusters, suggesting that they have younger stellar populations. The interpretation of the first result is complicated by the small number of clusters that have been studied so far, uncertainties in the initial mass function (IMF), and selection effects due to progenitor bias: if the morphologies of galaxies change with time, the sample of high-redshift early-type galaxies is only a subset of the sample of nearby early-type galaxies, leading to biased age estimates (see van Dokkum & Franx 2001). The interpretation of the second result is complicated by the fact that large corrections for selection effects need to be made (see Treu et al. 2005a, 2005b; van der Wel et al. 2005) and that galaxies are assumed to undergo no structural or dynamical changes with redshift.

The goal of the present paper is to better constrain the star formation epoch of massive field and cluster galaxies, using the FP.

<sup>1</sup> Based on observations obtained at the W. M. Keck Observatory, which is operated jointly by the California Institute of Technology and the University of California. Based on observations with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS 5-26555.

<sup>2</sup> Department of Astronomy, Yale University, New Haven, CT.

<sup>3</sup> Previous address: California Institute of Technology, Pasadena, CA.

<sup>4</sup> Space Telescope Science Institute, Baltimore, MD.

We present spatially resolved spectroscopic data from Keck and high-quality photometric data from the *Hubble Space Telescope* (*HST*) for a sample of early-type galaxies in three clusters at  $z \approx 0.5$ . The FPs of the three clusters are presented, discussed, and combined with literature data on the nearby Coma Cluster, a sample of nearby galaxies drawn from the Sloan Digital Sky Survey (SDSS), and 11 clusters at  $0.18 \leq z \leq 1.28$ . The evolution of the  $M/L$  ratio derived from this large sample is used to constrain the star formation epoch of massive cluster galaxies. We combine the cluster sample with recently published samples of field galaxies and determine the age difference between massive field and cluster galaxies using a self-consistent modeling approach that is different from previous studies. In two companion papers (van der Marel & van Dokkum 2006a, 2006b, hereafter vdMvD06a and vdMvD06b, respectively) we utilize our spatially resolved data to construct detailed dynamical models for the sample galaxies. The results are compared to those for local samples, providing an independent method of deriving  $M/L$  ratio evolution and allowing us to validate the many assumptions that enter into analyses based on the FP. The dynamical models also yield a normalized measure of the rotation rate [akin to  $(v/\sigma)^*$ ] of the sample galaxies, providing a method to identify S0 galaxies that have been visually misclassified as elliptical galaxies, based on their kinematics.

The structure of the present paper is as follows. The analysis of the spectroscopic and photometric data for the three  $z \approx 0.5$  clusters is described in §§ 2 and 3. In § 4 the FPs of the three clusters are presented and discussed. In § 5 results for the three clusters are combined with literature data on 11 additional distant clusters and two local samples, and the  $M/L_B$  evolution is determined from this large sample. Detailed information on the transformations of all of the literature data to our system is given in Appendices A and B. The measured evolution of the  $M/L_B$  ratio is interpreted in § 6. In this section the implications for the star formation epoch of massive cluster galaxies are discussed, and the cluster data are compared to previously published data for field early-type galaxies. The main results are summarized and discussed in § 7. We assume  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$  where needed.

## 2. SPECTROSCOPY

### 2.1. Cluster Selection

Galaxies were selected from the Smail et al. (1997) catalogs (MORPHS) of the  $z \approx 0.5$  galaxy clusters 3C 295, CL 0016+16, and CL 1601+42. These three clusters were selected based on their visibility at the time of our Keck observations and because they are among the most S0-deficient clusters in the MORPHS sample. The MORPHS sample itself was not selected according to strict criteria. Properties of the three clusters are listed in Table 1. They span a range of a factor of  $\sim 17$  in X-ray luminosity. 3C 295 and CL 1601+42 have very high measured velocity dispersions, possibly indicating substructure along the line of sight. The dominant galaxy in 3C 295 is a strong radio source. All MORPHS clusters have been observed with the *HST* WFPC2 camera, and bright galaxies in the WFPC2 fields were visually classified by the MORPHS team (Smail et al. 1997; Dressler et al. 1997).

### 2.2. Galaxy Selection and Observations

The sample selection was largely constrained by the geometry of the Keck Low Resolution Imager and Spectrograph (LRIS) masks. Priority was given to visually classified E and E/S0 galaxies with  $R_{702} < 21.5$  (3C 295 and CL 1601+42) or  $I_{814} < 20.8$  (CL 0016+16). For each cluster two masks were designed. The individual slits were tilted in order to optimally align them with

TABLE 1  
CLUSTER PROPERTIES

Parameter	3C 295	CL 0016+16	CL 1601+42
$z$ .....	0.460	0.546	0.538
$L_X^a$ ( $\times 10^{44} \text{ h}^{-2} \text{ ergs s}^{-1}$ ).....	3.20	5.88	0.35
$\sigma$ ( $\text{km s}^{-1}$ ).....	1670	1703	1166
F555W $T_{\text{exp}}$ (ks).....	...	12.6	...
F702W $T_{\text{exp}}$ (ks).....	12.6	...	16.8
F814W $T_{\text{exp}}$ (ks).....	...	16.8	...
E fraction <sup>b</sup> .....	0.51	0.58	0.33
S0 fraction <sup>b</sup> .....	0.21	0.15	0.12
S fraction <sup>b</sup> .....	0.28	0.27	0.55

NOTE.—Taken from Smail et al. (1997).

<sup>a</sup> 0.3–3.5 keV.

<sup>b</sup> Morphological fractions from Dressler et al. (1997).

the major axes of the galaxies, within the range allowed by geometrical restrictions. Some galaxies are contained in both masks; for those objects we used position angles on the sky that are offset by  $90^\circ$ , or the same position angle if the slit could be aligned with the major axis in both masks. A large, bright S0/Sb galaxy was included in one of the 3C 295 masks, to test whether rotation can be reliably measured for an “obvious” disk galaxy. Remaining space in the multislit masks was used to observe candidate lensed galaxies and random fainter objects in the WFPC2 field. These secondary “filler” objects are not discussed in this paper.

The three clusters were observed with LRIS on 2001 June 18–19. Conditions were variable, with part of the first night lost to clouds. The seeing varied from  $0.7''$  to  $0.9''$  during the run. For CL 0016+16 and CL 1601+42 exposures were obtained for both masks; for 3C 295 only one mask was observed. The red arm of LRIS was used with the  $900 \text{ line mm}^{-1}$  grating blazed at  $5500 \text{ \AA}$ . The  $1.1''$  wide slits provide a spectral resolution (as measured from sky emission lines)  $\sigma_{\text{instr}} \approx 65 \text{ km s}^{-1}$ . The observed wavelength range is different for each object (as it depends on its position in the multislit mask) but is typically  $3500\text{--}4800 \text{ \AA}$  in the rest frame. Total exposure times were 5400 s per mask for 3C 295 and CL 0016+16 and 9000 s per mask for CL 1601+42. The exposures were split into individual 1800 s exposures to facilitate cosmic-ray removal. The telescope was not moved in the slit direction between successive exposures, as this would have moved the objects out of the (tilted) slitlets. Calibration exposures (dome flats, arc lamps) were obtained in daytime. Dome flats were taken at a range of zenith angles.

### 2.3. Reduction

The reduction followed standard procedures for multislit spectroscopic data (see, e.g., Kelson et al. 2000b; van Dokkum & Stanford 2003). Each slitlet in the masks was treated as a separate long-slit spectrum. Bias was subtracted by fitting a low-order polynomial to the overscan regions. The CCD was read out using two amplifiers; the bias of each amplifier was fitted separately. Two bad columns were linearly interpolated. The dome flats were used to flat-field the data. Before flat-fielding, the dome flat exposures were divided by a polynomial fit in the  $\lambda$  and spatial directions, as the illumination of the CCD is different from that of the night-sky emission.

Bright cosmic rays were removed from each slitlet in the following way. A first-order approximation of the sky emission lines was subtracted by fitting a low-order polynomial. This step reduced the intensity of the lines but did not remove them completely because most slits are tilted. The sky line residuals and the

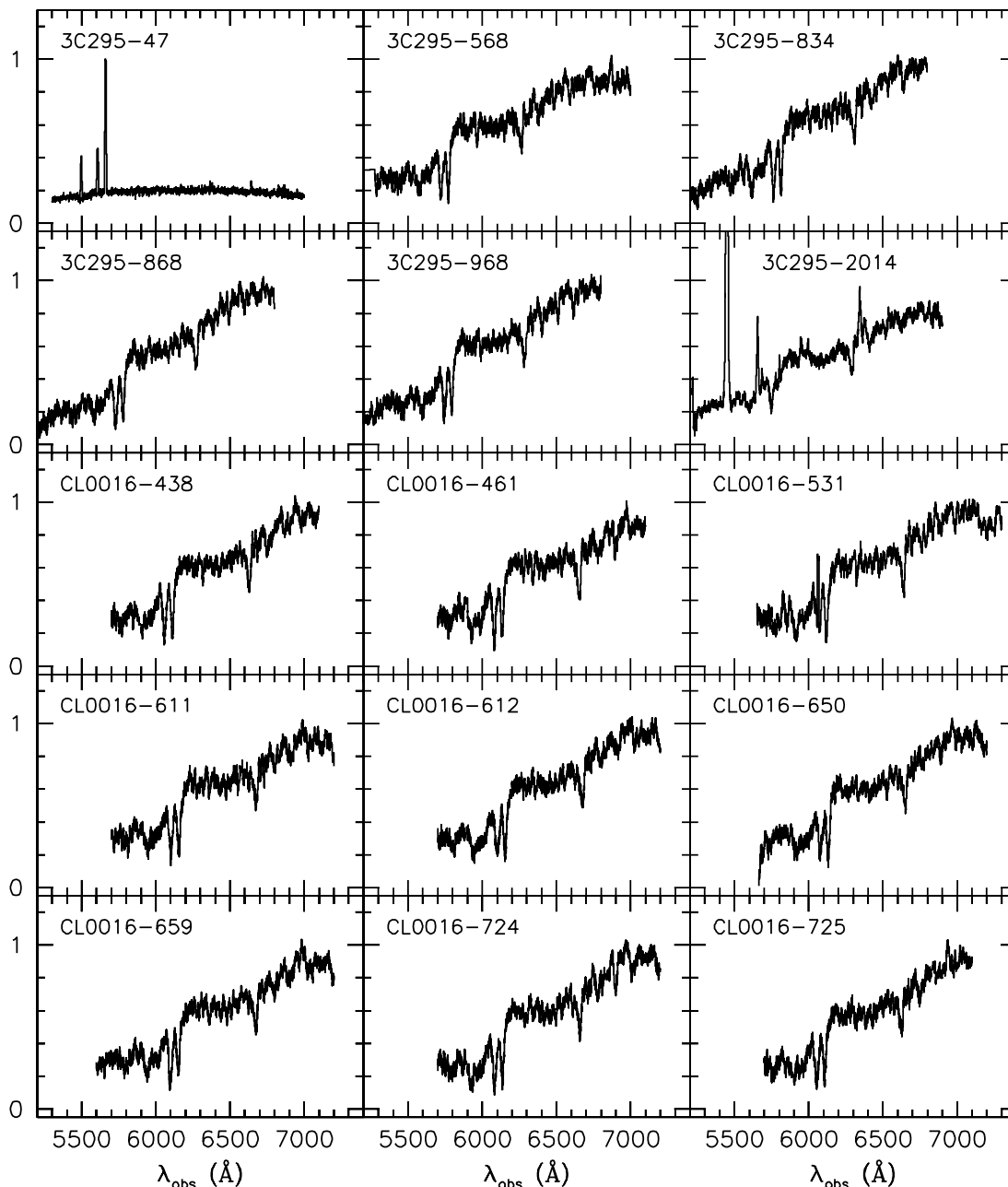


FIG. 1.—Keck spectra of all 27 targeted galaxies in the three cluster fields. The spectra are not binned or smoothed and have a resolution of  $\sigma \approx 65 \text{ km s}^{-1}$ . The vertical scale is arbitrary, as the spectra have not been calibrated. The spectroscopy shows that galaxies 3C 295–47 ( $z = 0.1308$ ) and CL 1601–270 ( $z = 0.5098$ ) are field galaxies unrelated to the clusters. These two galaxies are not included in the subsequent figures and analysis.

galaxy spectra were largely removed by subtracting the median of the exposures. Remaining residuals were removed by fitting a polynomial, first in the spatial direction and then in the wavelength direction. This three-step procedure proved to be much more effective in removing the light of galaxies and night-sky emission than simply subtracting the median. Cosmic rays were identified by comparing the counts to a noise model generated from the subtracted median image and the polynomial fits. Pixels deviating more than  $4\sigma$  were flagged as cosmic-ray hits.

Sky subtraction of tilted slits is difficult, as the standard fitting procedures require that sky lines are fairly well aligned with columns. Straightening the sky lines before subtracting them has the disadvantage of introducing aliasing effects due to the relatively large pixels of the LRIS CCD, which are very difficult to remove. More complex procedures have been developed that circumvent

these problems (e.g., Kelson 2003); however, as we are working in a wavelength region mostly blueward of the ubiquitous OH lines and the galaxies are relatively bright, we found that an iterative approach provided satisfactory results. At this point in the reduction we subtracted sky lines by simply fitting a low-order polynomial in the spatial direction, masking the galaxy spectrum and any other objects in the slit. While not completely removing the lines, this step greatly reduces the aliasing effects that occur when resampling the lines after wavelength calibration.

Wavelength calibration was done using arc lamp exposures, with all lamps on (Hg, Ne, Ar, Cd, and Zn). Line identification was done separately for each row, as the solution is a strong function of position for these tilted slits. The O I  $\lambda 6300.4$  sky line was used to apply small corrections to the zero points of the wavelength solutions. The S-distortion of the spectra was determined

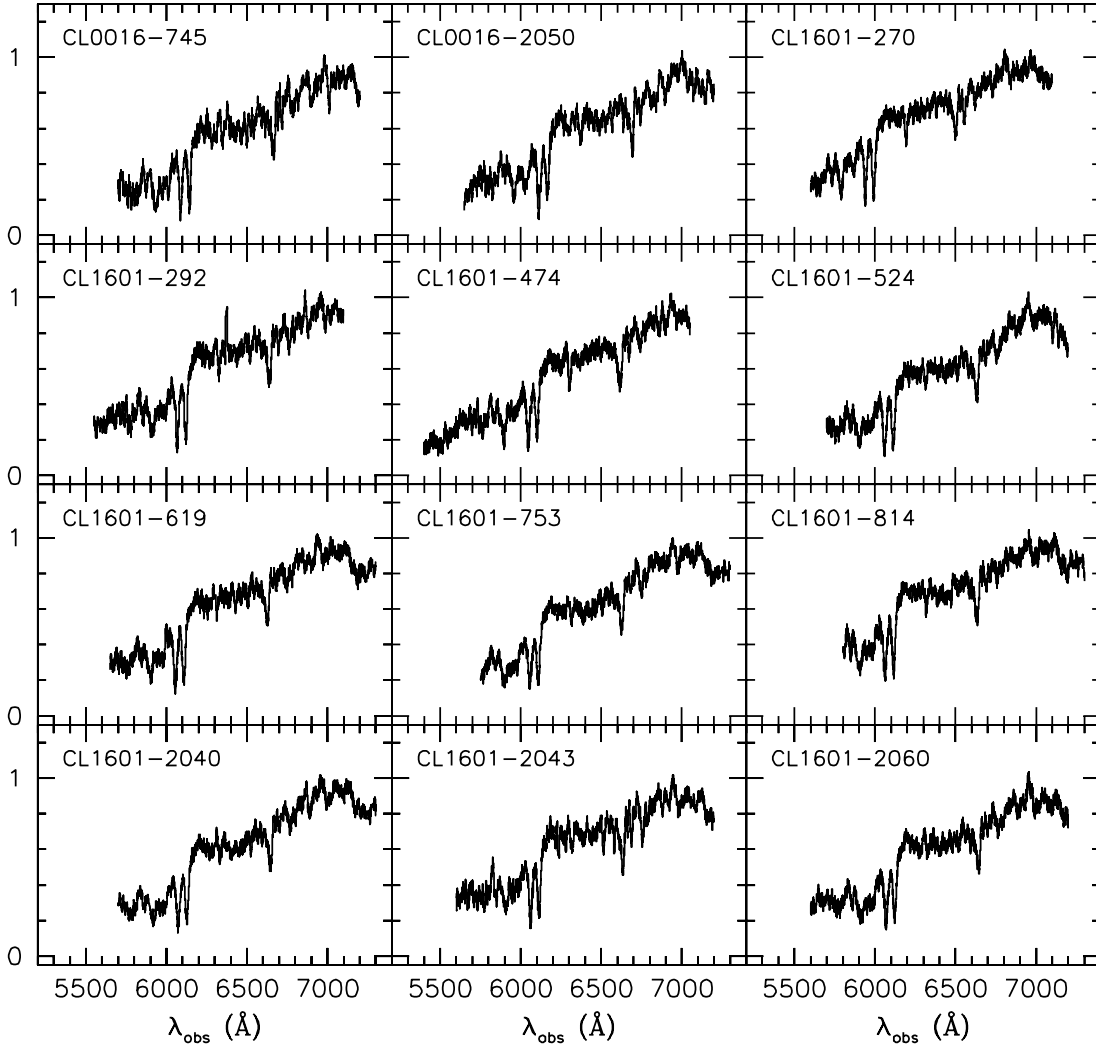


FIG. 1—Continued

by fitting the wavelength dependence of the position of the galaxy spectra. Two-dimensional polynomials were fitted to the measured positions of the sky lines and galaxy spectra, and these were used to rectify the spectra. In this procedure the spectra are re-sampled only once. After rectification, remaining sky line residuals were removed by fitting a polynomial (masking the galaxy spectra), and remaining faint cosmic rays were removed with L.A.COSMIC (van Dokkum 2001).

Spectra of the 27 target objects are shown in Figure 1. They were created by averaging the central five rows ( $1.1''$ ) and have not been weighted, smoothed, or binned. The sampling is  $38 \text{ km s}^{-1} \text{ pixel}^{-1}$ , and  $\sigma_{\text{instr}} \approx 65 \text{ km s}^{-1}$ . The signal-to-noise ratio (S/N) varies from  $\sim 20$  to  $\sim 60 \text{ Å}^{-1}$ . Note that the spectra are not flux calibrated.

#### 2.4. Redshifts

Redshifts were determined using the cross-correlation program XCSAO in IRAF or, in one case, from emission lines. A nearby early-type galaxy was used as the template in the cross-correlation. The redshifts are listed in Table 2 along with the Smail et al. (1997) morphological classifications. Of 27 observed galaxies, 25 are at the redshift of the cluster. Of those 25, 24 have an early-type spectrum without strong emission lines. The exception is 3C 295–2014, which shows broad emission lines super-

posed on an early-type spectrum. The presence of active galactic nucleus (AGN) features in this object is not surprising as it hosts the bright radio source 3C 295, the namesake of the cluster.

The two galaxies with deviant redshifts are 3C 295–47 and CL 1601–270. 3C 295–47 is an emission-line object at  $z = 0.1308$ . The lines are strong and narrow; the  $[\text{O III}] \lambda 5007$  line has a rest-frame equivalent width  $W_\lambda \approx 45 \text{ Å}$  and is unresolved at our resolution of  $\approx 65 \text{ km s}^{-1}$ . The galaxy resembles the compact narrow emission line galaxies described by Guzman et al. (1996). CL 1601–270 is a field early-type galaxy  $\sim 10^4 \text{ km s}^{-1}$  removed from the mean recession velocity of CL 1601+42. These galaxies are not included in the subsequent figures and analysis. We note that galaxy CL 1601–270 is included in the modeling of vdMvD06a and vdMvD06b.

#### 2.5. Velocity Dispersions

Internal velocity dispersions were determined from direct fits of the galaxy spectra to template star spectra. The fitting methodology is explained in detail elsewhere (see, e.g., van Dokkum & Franx 1996; Kelson et al. 2000b; van Dokkum & Stanford 2003). Briefly, the template stars were broadened such that their spectral resolution matches the instrumental resolution of the observations. Residuals of night-sky lines, Balmer absorption lines, and

TABLE 2  
GALAXY SAMPLE

ID <sup>a</sup>	Type <sup>a</sup>	$z$	$\sigma^b$ (km s <sup>-1</sup> )	$\log r_e^c$ (arcsec)	$\mu_B^d$	$V_{555} - I_{814}$	Remarks
3C 295–47 .....	E/S0	0.1308	...	...	...	...	Field galaxy <sup>c</sup>
3C 295–568 .....	S0/Sb	0.4545	152 ± 10	−0.092	23.68	...	Spiral galaxy <sup>c</sup>
3C 295–834 .....	E	0.4651	165 ± 14	−0.801	21.35	...	
3C 295–868 .....	S0	0.4565	197 ± 18	−0.439	22.93	...	
3C 295–968 .....	E	0.4601	213 ± 12	−0.176	23.64	...	
3C 295–2014 .....	E/S0	0.4616	284 ± 22	0.583	25.33	...	AGN (3C 295)
CL 0016–438 .....	E	0.5399	229 ± 14	−0.269	22.96	2.460	
CL 0016–461 .....	E	0.5458	268 ± 24	−0.532	22.37	2.461	
CL 0016–531 .....	E	0.5420	209 ± 21	−0.687	21.98	2.443	
CL 0016–611 .....	E	0.5509	169 ± 18	−0.441	22.71	2.466	
CL 0016–612 .....	E	0.5508	279 ± 39	−0.322	22.75	2.574	
CL 0016–650 .....	E/S0	0.5445	172 ± 20	0.633	26.35	2.482	
CL 0016–659 .....	E	0.5502	264 ± 27	−0.293	23.13	2.441	
CL 0016–724 .....	E	0.5463	252 ± 17	0.462	25.46	2.582	
CL 0016–725 .....	E	0.5389	226 ± 18	0.739	26.18	2.531	
CL 0016–745 .....	E	0.5478	124 ± 14	−0.105	24.06	2.519	
CL 0016–2050 .....	E	0.5536	125 ± 18	−0.444	23.11	2.366	
CL 1601–270 .....	S0/E	0.5098	186 ± 12	−0.315	22.84	...	Field early-type <sup>c</sup>
CL 1601–292 .....	E	0.5425	179 ± 12	−0.301	22.99	...	
CL 1601–474 .....	E/S0	0.5376	313 ± 51	0.033	24.30	...	Merger <sup>c</sup>
CL 1601–524 .....	E	0.5403	226 ± 16	−0.199	23.50	...	
CL 1601–619 .....	E	0.5392	232 ± 12	−0.044	24.00	...	
CL 1601–753 .....	E	0.5395	281 ± 12	0.131	24.20	...	
CL 1601–814 .....	E	0.5409	226 ± 19	−0.452	22.88	...	
CL 1601–2040 .....	E	0.5439	256 ± 10	0.062	23.83	...	
CL 1601–2043 .....	E	0.5407	163 ± 14	−0.461	23.13	...	
CL 1601–2060 .....	E	0.5430	235 ± 36	−0.165	23.43	...	

<sup>a</sup> From Smail et al. (1997).

<sup>b</sup> Velocity dispersion corrected to a 3.4'' diameter aperture at the distance of the Coma Cluster.

<sup>c</sup> Circularized effective radius.

<sup>d</sup> Surface brightness at the effective radius, in rest-frame  $B$ , corrected for Galactic extinction but not corrected for cosmological surface brightness dimming.

<sup>e</sup> Not included in the analysis of the  $M/L$  evolution.

(in the case of 3C 295–2014) emission lines were masked in the fits. The wavelength region used in the fit was 4100–4650 Å; changing the fitting region has a negligible effect on the measured dispersions. For consistency with previous work (e.g., van Dokkum & Franx 1996) a K0 giant star was adopted as the template star. Varying the template star and the continuum filtering produces  $\lesssim 5\%$  variations in the dispersions.

For each galaxy (except 3C 295–47) velocity dispersions were measured from the unweighted average spectrum of the five central rows, corresponding to a  $1.1'' \times 1.1''$  square aperture. For consistency with previous work the measured dispersions were corrected to the equivalent of a 3.4'' diameter aperture at the distance of the Coma Cluster, using the empirical logarithmic correction given by Jørgensen et al. (1995b). For  $z \approx 0.5$  the correction is approximately 7%. Corrected dispersions are listed in Table 2. The median random error is 8%, and in every case the error is less than 15%.

## 2.6. Spatially Resolved Kinematics

In vdMvD06a we present dynamical models of the galaxies, derived from spatially resolved kinematics and photometry. Among other parameters, these models provide  $M/L$  ratios of the galaxies free of many of the assumptions that enter the FP analysis. The radial profiles of velocity and velocity dispersion were determined in the following way. Velocity dispersion measurements were made at seven positions along the slit: the average of rows −7, −6, −5; the average of rows −4, −3, −2; row −1; the central row 0; row +1; the average of rows 2, 3, 4; and the average of rows 5, 6, 7. As the pixel

size is  $0.215''$ , these positions correspond to  $-1.29''$ ,  $-0.64''$ ,  $-0.22''$ ,  $0''$ ,  $0.22''$ ,  $0.64''$ , and  $1.29''$  from the center of the galaxy. Radial velocities were measured using XCSAO, as in some cases velocities can still be measured when the S/N is too low to measure dispersions. Dispersions were deemed unreliable when the random error exceeds 20%; velocities were discarded when the Tonry & Davis (1979)  $R$ -value is lower than 3. The resulting profiles are shown in vdMvD06a.

## 3. PHOTOMETRY

### 3.1. WFPC2 Images

The three clusters were observed with WFPC2 on *HST* as part of the MORPHS cluster program. Details of the observations and data characteristics are given in Smail et al. (1997). 3C 295 and CL 1601+42 were observed in one band ( $R_{702}$ ) only; CL 0016+16 was observed in two filters ( $V_{555}$  and  $I_{814}$ ). Total exposure times are listed in Table 1.

We obtained the raw data from the *HST* archive. For each cluster, the exposure time was divided over two positions offset by  $\approx 2''$ . The shifts are not integer numbers of pixels, and rather than combining the two positions, we reduced each pointing separately. This procedure has the advantages that no interpolation is required and that we can assess the uncertainties in derived parameters from two independent data sets, with different subpixel sampling. The IRAF *crrej* task was used to combine individual frames at each position. Remaining cosmic rays and hot pixels were removed using L.A.COSMIC.

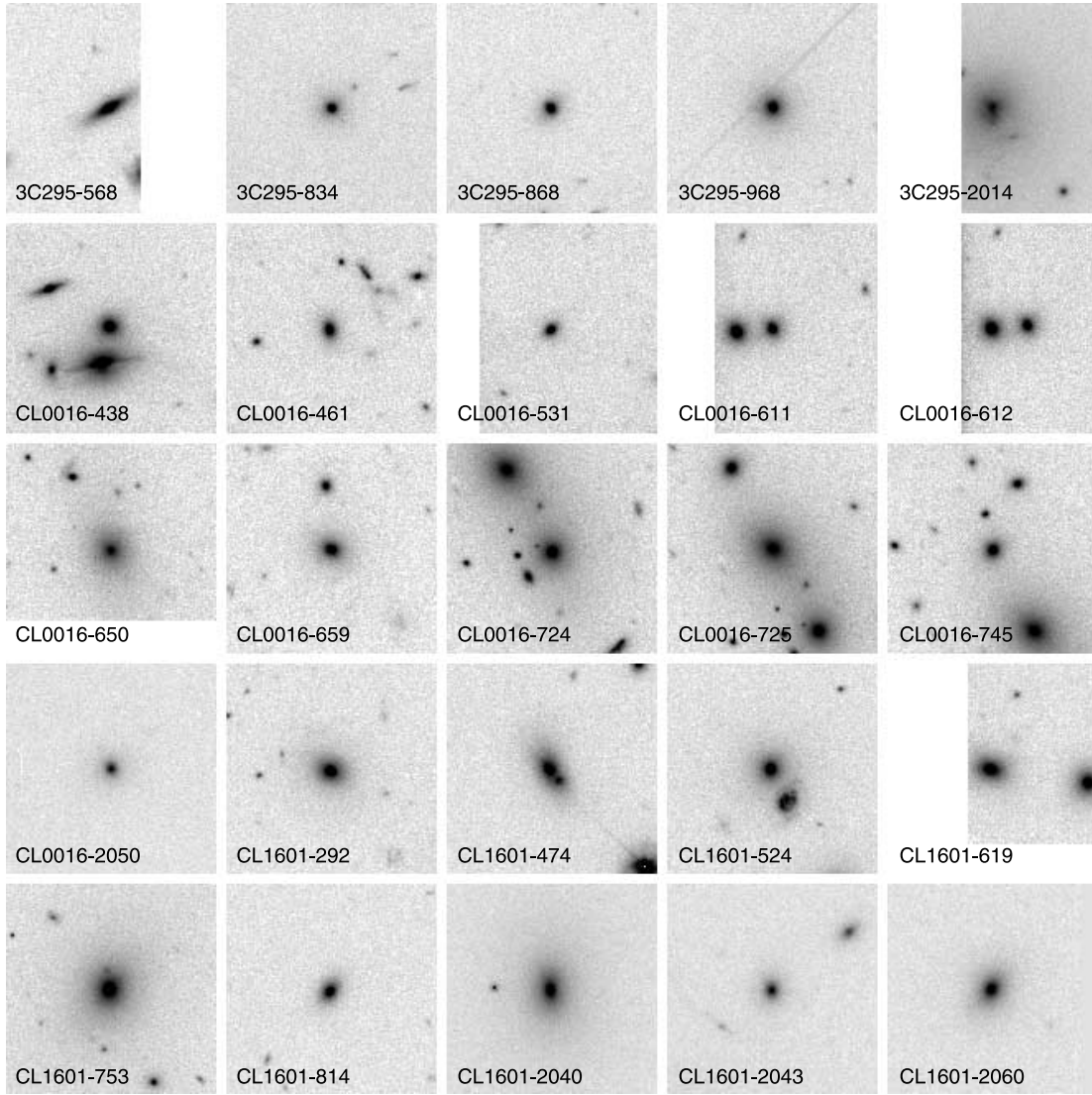


FIG. 2.—*HST* WFPC2 images of the 25 cluster early-type galaxies observed spectroscopically.

Images of the 25 cluster galaxies are shown in Figure 2. As expected, most of the galaxies have very regular morphologies with no obvious spiral arms or other fine structure. There are three notable exceptions: 3C 295–568 is the S0/Sb galaxy added to test our ability to measure rotation curves; 3C 295–2014 (the central cluster galaxy) has structure in its core, very likely associated with its active nucleus; and CL 1601–474 consists of two overlapping galaxies. Both components of CL 1601–474 show morphological disturbances, strongly suggesting that this is a merger in progress. The object is flagged as “interaction/merger?” in the Smail et al. (1997) catalog. Interestingly, the galaxy has an early-type spectrum, similar to the red mergers in the  $z = 0.83$  cluster MS 1054–03 (Tran et al. 2005a).

### 3.2. Structural Parameters

#### 3.2.1. Fitting

The galaxy images were fitted with two-dimensional models, created by convolving  $r^{1/4}$  law light distributions with point-spread functions (PSFs). A separate PSF was used for each object, as the WFPC2 PSF depends on the position of the object on the chip. The PSFs were generated with the Tiny Tim 6.1a software

(Krist 1995). The fitting procedure is described in van Dokkum & Franx (1996). Fit parameters in the  $\chi^2$  minimalization are the  $x, y$  position, the effective radius  $r_e$ , the surface brightness at the effective radius  $I_e$ , the position angle, the ellipticity, and the sky value. Nearby objects and the edges of CCDs were masked in the fits. We inspected the residual images, created by subtracting the best-fitting models from the data, to assess the quality of the fit. In most cases the model provides a good fit to the data; exceptions are the S0/Sb galaxy 3C 295–568 and the active nucleus 3C 295–2014.

We empirically determined the errors in the fitted parameters by comparing the values derived from the two independent pointings, as the formal errors from the  $\chi^2$  fit do not include the effects of undersampling. The results are shown in Figure 3. The uncertainties in  $r_e$  and  $I_e$  are highly correlated, and as is well known, the correlation is almost parallel to the relation  $I_e \propto r_e^{-1.2}$  (solid line) that enters the FP (see, e.g., Jørgensen et al. 1995a; van Dokkum & Franx 1996; Kelson et al. 2000a). Not surprisingly, the correlation is also well described by the expected relation for total flux conservation ( $I_e \propto r_e^{-2}$ ; dashed line). The rms uncertainty in a single measurement is 0.012 in  $\log r_e$ , 0.018 in  $\log I_e$ , and 0.003 in the FP combination  $\log r_e + 0.8 \log I_e$ . These uncertainties are

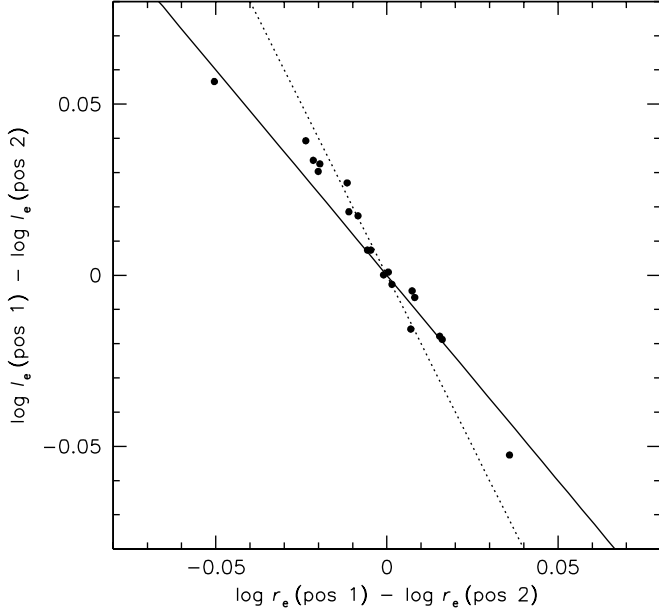


FIG. 3.—Comparison of effective radii  $r_e$  and surface brightnesses  $I_e$  measured from two independent *HST* pointings. The errors in the two parameters are highly correlated. The dashed line is the relation  $I_e \propto r_e^{-2}$ , and the solid line is the relation  $I_e \propto r_e^{-1.2}$  that enters the FP.

negligible compared to the random errors in the velocity dispersions and also to the systematic errors in the photometric calibration (see § 3.2.2).

### 3.2.2. Calibration

Measured effective radii were converted to arcseconds using a pixel scale of  $0.09993'' \text{ pixel}^{-1}$  for the Wide Field camera and  $0.04555'' \text{ pixels}^{-1}$  for the Planetary Camera. They were converted to kiloparsecs using our adopted cosmology. We note that the analysis in § 5 depends on the value of  $\Omega$  but not on the Hubble constant (see, e.g., van Dokkum et al. 1998a).

Surface brightnesses were converted to the *HST* Vega system using the appropriate zero points for each of the WFPC2 chips, as obtained from the WFPC2 Data Handbook,<sup>5</sup> and corrected for Galactic extinction using the Schlegel et al. (1998) maps. The extinction is highest in the CL 0016+16 field ( $\approx 0.11$  mag in the  $I_{814}$  band). We did not correct for the long versus short anomaly, as our fields are not very crowded, or the position-dependent charge transfer efficiency (CTE), as the backgrounds are relatively high.

For a meaningful comparison of galaxies observed at different redshifts the photometry has to be transformed to a common rest-frame band. At  $z \sim 0.5$  the observed  $R_{702}$  band is close to the rest-frame  $B$  band (denoted  $B_z$ ). Following the methodology of van Dokkum & Franx (1996), we derived the following transformations from observed to rest-frame magnitudes:

$$B_z = R_{702} + 0.23(V_{555} - R_{702}) + 0.74, \quad (1)$$

$$B_z = I_{814} + 0.30(V_{555} - I_{814}) + 0.89, \quad (2)$$

$$B_z = R_{702} + 0.05(V_{555} - R_{702}) + 0.85, \quad (3)$$

for 3C 295, CL 0016+16, and CL 1601+42, respectively. For the synthetic  $B$  band we used the Bessell (1990)  $BX$  filter. This filter is very similar to the Buser & Kurucz (1978)  $B_2$  filter, which

is used by Bruzual & Charlot (2003) for their stellar population synthesis calculations.

Using templates from Coleman et al. (1980), we find that the transformations are independent of spectral type to  $\approx 0.03$ . However, the colors that enter the transformations are obviously a strong function of spectral type. Measured colors are only available for CL 0016+16 (Smail et al. 1997), as the other clusters were observed in only one filter. For 3C 295 and CL 1601+42 we used synthetic colors derived from a redshifted Coleman et al. (1980) E/S0 template, as the measured colors for CL 0016+16 are a good match to this template. We find  $V_{555} - R_{702} = 1.71$  for 3C 295 and  $V_{555} - R_{702} = 1.81$  for CL 1601+42.

The uncertainties in the calibration are a combination of uncertainties in the WFPC2 zero points ( $\approx 0.03$  mag), the long-short anomaly and CTE effects ( $\approx 0.04$  mag), the extinction ( $\approx 0.02$  mag), and the conversion to rest-frame  $B$  magnitudes. The uncertainty introduced by the lack of measured colors is only a few percent for CL 1601+42, as the color term is small. However, the effect can be substantial for the lower redshift cluster 3C 295: using an Sb/c template rather than an E/S0 template gives  $B_z$  magnitudes that are  $\approx 0.15$  mag brighter. Hence, we estimate that the typical uncertainty is  $\approx 0.10$  mag for this cluster. We assume that the systematic errors are uncorrelated and can be added in quadrature (see, e.g., Freedman et al. 2001). The combined calibration errors in the surface brightnesses are  $\approx 0.11$  mag for 3C 295 and  $\approx 0.06$  mag for CL 0016+16 and CL 1601+42.

## 4. THE FUNDAMENTAL PLANE

The FP is an empirical relation between the effective radii  $r_e$ , surface brightness at the effective radii  $I_e$ , and the velocity dispersions  $\sigma$  of early-type galaxies of the form (Djorgovski & Davis 1987)

$$\log r_e = a \log \sigma + b \log I_e + c. \quad (4)$$

The coefficients  $a$  and  $b$  depend on wavelength, sample selection, and fitting method (see, e.g., Bernardi et al. 2003b). Jørgensen et al. (1996) find  $a = 1.20 \pm 0.06$  and  $b = -0.83 \pm 0.02$  in the  $B$  band. We define  $\log I_e \equiv \mu_e / -2.5$ , with  $\mu_e$  the rest-frame  $B$ -band surface brightness at the effective radius in  $\text{mag arcsec}^{-2}$ , corrected for  $(1+z)^4$  cosmological surface brightness dimming. We note that for an  $r^{1/4}$  law fit the relation between  $\mu_e$  and  $\langle \mu \rangle_e$ , the average surface brightness within the effective radius, is given by  $\mu_e = \langle \mu \rangle_e + 1.393$ .

### 4.1. Edge-on Projection

The edge-on projections of the FPs in the three clusters are shown in Figure 4. Filled circles are E and E/S0 galaxies, the filled triangle is an S0 galaxy, the open square is a merger system, and the open circle is the S0/Sb galaxy that was added to test whether we can measure rotation curves. Small dots are galaxies in the nearby Coma Cluster from Jørgensen et al. (1996) (see § A1).

The FP is well defined in each cluster: the elliptical galaxies in the  $z \approx 0.5$  clusters qualitatively follow very similar relations as the Coma galaxies. The FPs are offset with respect to Coma because of passive luminosity evolution of the stellar populations in the galaxies (see, e.g., van Dokkum & Franx 1996). The S0/Sb galaxy 3C 295–568 and the merger CL 1601–474 are outliers. The spiral galaxy falls below the relation defined by the other galaxies in 3C 295, consistent with results for spiral galaxies in the  $z = 0.33$  cluster CL 1358+62 (Kelson et al. 2000c). The merging galaxy falls above the CL 1601+42 relation, possibly because its velocity dispersion is overestimated due to the presence of light from its companion in the slit. Both galaxies were excluded from the analysis in § 5.

<sup>5</sup> See [http://www.stsci.edu/instruments/wfpc2/Wfpc2\\_dhb/WFPC2\\_longdnhbcover.html](http://www.stsci.edu/instruments/wfpc2/Wfpc2_dhb/WFPC2_longdnhbcover.html).

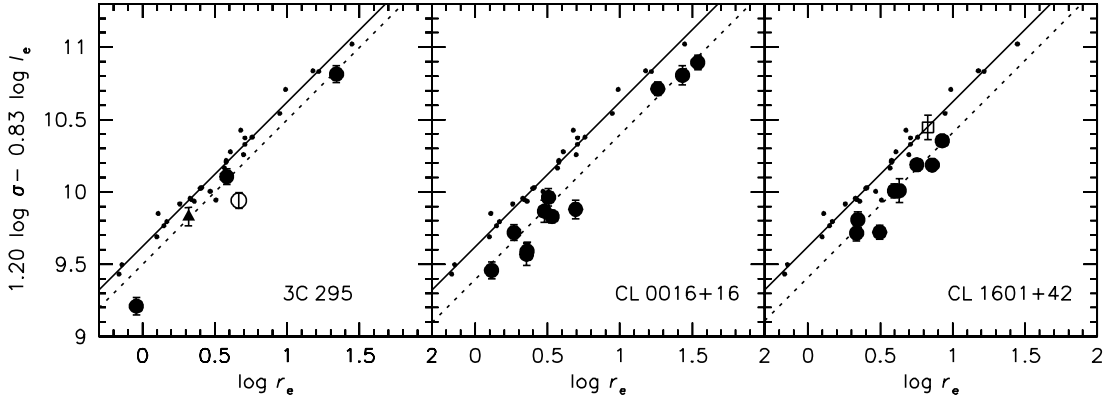


FIG. 4.—Edge-on projection of the FP in the three clusters. Filled circles are E and E/S0 galaxies, the filled triangle is the S0 galaxy 3C 295–868, the open circle is the S0/Sb galaxy 3C 295–568, and the open square is the merger CL 1601–474. Dots are galaxies in the nearby Coma Cluster. The solid line is a fit to Coma, and the dashed lines show the best-fitting offsets from the Coma FP for each cluster.

#### 4.2. Scatter and Tilt of the Fundamental Plane

The galaxy samples in the individual clusters are too small to reliably determine the scatter in the FP. Hence, the data for the three clusters are combined. For each galaxy the residual from the FP was calculated using

$$\Delta_{\text{FP}}(\log r_e) = \log r_e - (1.20 \log \sigma - 0.83 \log I_e + c) - 0.461(z - 0.024). \quad (5)$$

This procedure assumes that the intrinsic cluster-to-cluster scatter in the zero point of the FP is small compared to the galaxy-to-galaxy scatter within the clusters. The redshift term compensates for differential luminosity evolution over the redshift range spanned by the clusters, and the zero point  $c = -9.626$  is that of the  $z = 0.024$  Coma Cluster (see § 5).

The scatter is calculated using the biweight estimator and its uncertainty is determined from bootstrap resampling. For the full sample of E and E/S0 galaxies (excluding the merger CL 1601–474) the scatter is  $0.095 \pm 0.028$  in  $\log r_e$ . This scatter is slightly higher than that measured in local clusters; e.g., Jørgensen et al. (1996) find an rms scatter of 0.071 in  $\log r_e$ . We use Monte Carlo simulations to determine whether the measured scatter can be fully explained by measurement errors (which are dominated by the errors in  $\log \sigma$ ). The errors imply an expected scatter of  $0.048 \pm 0.009$ , and we conclude that the intrinsic scatter in this sample is  $0.082 \pm 0.028$  in  $\log r_e$ .

We investigate the cause of this scatter in Figure 5, which shows the relation of the residuals from the FP with effective radius, velocity dispersion, and mass. Masses were determined using

$$\log M = 2 \log \sigma + \log r_e + 6.07. \quad (6)$$

The masses are approximate, as the constant in equation (6) depends on the details of the galaxy model, such as the brightness profile and the internal dynamical structure. There are clear systematic trends in the  $z \approx 0.5$  sample that are not present in the Coma sample. The residuals correlate with velocity dispersion and with mass: a Spearman rank test gives a probability of 99.8% that  $\Delta \log r_e$  and  $\sigma$  are correlated and a probability of 96% that  $\Delta \log r_e$  and  $M$  are correlated. Correlations with  $\log r_e$  and  $\log I_e$  are not significant according to the Spearman rank test. When the sample is limited to galaxies having  $\sigma > 200 \text{ km s}^{-1}$ , the observed scatter in  $\Delta \log r_e$  reduces to  $0.055 \pm 0.019$ ; galaxies with  $M > 10^{11} M_\odot$  give a scatter of  $0.064 \pm 0.020$ . Both values are consistent with the scatter due to measurement errors alone, and we conclude that the intrinsic scatter in the full sample of E/S0 galaxies is caused by the galaxies with the lowest masses.

Similar correlations were found by Wuyts et al. (2004) for a cluster at  $z = 0.58$  and by Jørgensen et al. (2006) for two clusters at  $z = 0.83$  and  $0.89$ . Furthermore, recent studies of field early-type galaxies at  $0.5 < z < 1.1$  have found the same qualitative trends (Treu et al. 2005b; van der Wel et al. 2005; di Serego Alighieri et al. 2005). The usual interpretation of these trends is

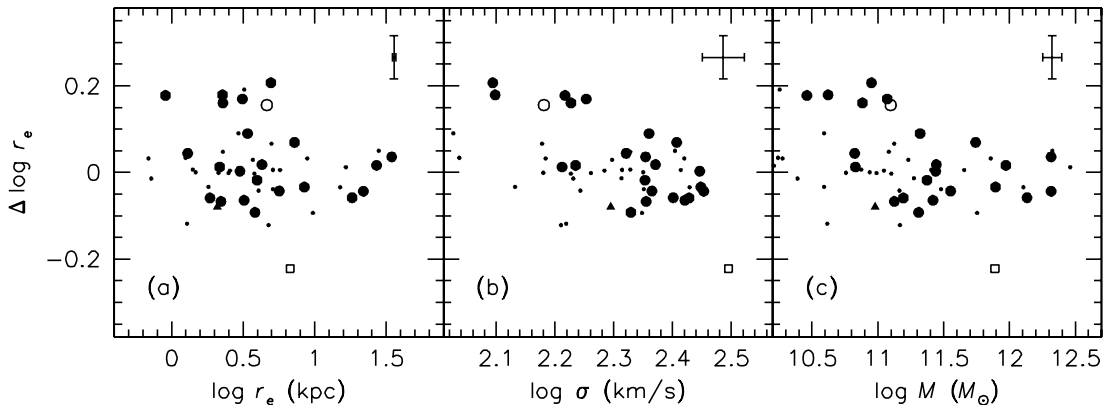


FIG. 5.—Correlation of the residual from the FP with (a) effective radius, (b) velocity dispersion, and (c) mass, for galaxies in the three clusters. Symbols have the same meaning as in Fig. 4. Galaxies with low masses or low velocity dispersions appear to systematically deviate with respect to the Coma Cluster sample.



that they are due to a combination of selection effects and evolution in the “tilt” of the FP (see, e.g., Treu et al. 2005a). The evolution in the tilt is taken as evidence for differential evolution in the  $M/L$  ratio for galaxies of different mass, resulting from an underlying age-mass relation (with low-mass early-type galaxies being younger than high-mass early-type galaxies; see, e.g., Treu et al. 2005a; van der Wel et al. 2005). In vdMvD06b we show that these observed trends may be in part due to structural or kinematic evolution of low-mass galaxies, as opposed to evolution of their  $M/L$  ratios. In the following we restrict the discussion to galaxies with masses  $>10^{11} M_{\odot}$ . As shown in vdMvD06b, FP evolution is an unbiased measure of  $M/L$  evolution for galaxies in this mass range.

Finally, we note that Kelson et al. (2000c) and Moran et al. (2005) do not find evidence for a change in the tilt of the FP for the clusters CL 1358+62 at  $z = 0.33$  and CL 0024+16 at  $z = 0.39$ , respectively, despite the use of large, high-quality data sets. We also note that Wuyts et al. (2004) and Moran et al. (2005) report an increased intrinsic scatter in the FP compared to nearby clusters even for high-mass galaxies, which is not seen in our sample or in CL 1358+62 (Kelson et al. 2000c). The available evidence suggests that the FPs of intermediate-redshift clusters show a range of properties when studied in detail, perhaps due to the stochastic nature of the growth of clusters (see, e.g., Tran et al. 2005b; Moran et al. 2005). Studies of individual clusters should therefore be interpreted with caution.

## 5. EVOLUTION OF THE MEAN $M/L$ RATIO

### 5.1. Procedure

Evolution of the zero point of the FP can be interpreted as a systematic change of the mean  $M/L$  ratio with redshift. The conversion of zero-point offsets to offsets in  $M/L$  ratio assumes that early-type galaxies form a homologous family and that the FP is a manifestation of an underlying relation between the  $M/L$  ratio of galaxies and other parameters (e.g., Faber et al. 1987). Starting from the empirical FP relation and assuming  $M \propto \sigma^2 r_e$  and  $L \propto I_e r_e^2$ , this underlying relation is

$$M/L \propto \sigma^{2+a/b} r_e^{-(1+b)/b} \quad (7)$$

(e.g., Kelson et al. 2000c). The Jørgensen et al. (1996)  $B$ -band coefficients imply that  $M/L$  is largely a function of mass:  $M/L \propto M^{0.28} r_e^{-0.08}$ . Different coefficients imply different relations; e.g., using a large sample of galaxies drawn from the SDSS, Bernardi et al. (2003b) find that  $M/L$  correlates with effective radius with little or no dependence on  $\sigma$ .

Rather than determine relations between  $M/L$  and other observables, we follow the usual practice of determining the evolution in  $M/L$  directly from the fundamental plane offset with respect to  $z \approx 0$ :

$$\Delta \log(M/L)_z = \log(M/L)_z - \log(M/L)_0 = (c_z - c_0)/b. \quad (8)$$

This procedure assumes that the observed change in intercept of the FP is caused by evolution of the  $M/L$  ratio and that the coefficients  $a$  and  $b$  do not depend on the redshift. In vdMvD06b it is demonstrated that FP evolution thus defined yields an unbiased measure of  $M/L$  evolution for galaxies with velocity dispersions  $\gtrsim 200 \text{ km s}^{-1}$  or masses  $\gtrsim 10^{11} M_{\odot}$ . In practice, we use the Coma Cluster of galaxies to define the zero point  $c_0$  (see Appendix A). This choice is arbitrary, as adding a constant to all values of  $\Delta \log(M/L)_z$  does not change the results of our subsequent analysis in any way.

The coefficient  $c_z$  for each cluster is determined by calculating the residual from the FP for each early-type galaxy with  $M > 10^{11} M_{\odot}$ :

$$c_{z,i} = \log r_e - (a \log \sigma + b \log I_e). \quad (9)$$

The cluster offset  $c_z$  is the biweight mean (Beers et al. 1990) of the distribution of  $c_{z,i}$ . For consistency with Jørgensen et al. (1996) we use  $a = 1.20$  and  $b = -0.83$  in equations (8) and (9). As we show later the results are only very weakly dependent on the values of these parameters. For each cluster a random error and a systematic error are determined. The random error is the formal uncertainty in the biweight mean, or the expected uncertainty due to the errors in the individual measurements, if the latter exceeds the former. As discussed in § 3.2.1, random errors in the velocity dispersions dominate over those in the structural parameters. The minimum random error in  $c_z$  is therefore approximately  $\overline{\delta\sigma}/\sqrt{n}$ , with  $\overline{\delta\sigma}$  the average random uncertainty in the velocity dispersion and  $n$  the number of galaxies in the sample. Systematic errors are described for each cluster individually below.

### 5.2. Other Clusters

Including the three clusters presented here, there are now more than a dozen clusters in the redshift range 0.2–1.3 for which FP measurements have been made. Furthermore, at low redshift the SDSS has provided important new information unhampered by uncertainties in peculiar motions and small number statistics. We combine published data with our own to compile a sample consisting of the Coma Cluster, the SDSS sample, and 14 distant clusters with homogeneous FP measurements. With one exception, the criteria for including distant clusters are (1) published data for individual galaxies and (2) sufficient information to bring the data to a consistent system (i.e., colors, aperture corrections, etc.). The exception is RX J1226+33 at  $z = 0.892$ : data of individual galaxies in that cluster are not yet published, but they were provided to us by I. Jørgensen. Details on the individual clusters and the derivation of systematic uncertainties are given in Appendix A, and the offsets are listed in Table 3. The uncertainties include neither the propagation of uncertainties in  $\Omega_m$  and  $\Omega_{\Lambda}$  nor the propagation of uncertainties in distance due to cosmic variance in  $H_0$  (i.e., the expansion factor having somewhat different value when averaged over small scales than when averaged over large scales). This simplification is justified by the fact that these uncertainties are generally not the dominant uncertainty in the analysis (see vdMvD06b).

### 5.3. Observed Evolution and Scatter

The observed evolution of the  $M/L_B$  ratio of massive cluster galaxies is shown in Figure 6. There is a clear relation, with galaxies in clusters at high redshift having lower  $M/L_B$  ratios than those at low redshift, consistent with previous studies of smaller samples (e.g., van Dokkum & Franx 1996; van Dokkum et al. 1998a; Kelson et al. 1997; Holden et al. 2005). The best-fitting linear function has a slope

$$d \log(M/L_B)/dz = -0.555 \pm 0.042 \quad (10)$$

and is indicated in Figure 6. This evolution is stronger than inferred from previous studies that were based on smaller samples. For example, van Dokkum & Stanford (2003) find a slope of  $-0.460 \pm 0.039$ , and Holden et al. (2005) find  $-0.426 \pm 0.026$ . This difference is largely due to sample size: if we limit the sample to the six clusters from van Dokkum & Stanford (2003) or the seven clusters from Holden et al. (2005), we find

TABLE 3  
 $M/L_B$  OFFSETS

Sample	$z$	$N$	$\Delta \log (M/L_B)$
Coma .....	0.024	16	$0.000 \pm 0.029$
SDSS, $n_{\text{den}} > 20$ .....	0.109	171	$+0.021 \pm 0.031$
Abell 2218 .....	0.176	8	$+0.009 \pm 0.037$
Abell 665 .....	0.183	5	$-0.006 \pm 0.040$
Abell 2390 .....	0.228	5	$-0.035 \pm 0.065$
CL 1358+62 .....	0.327	16	$-0.162 \pm 0.029$
CL 0024+16 .....	0.391	6	$-0.160 \pm 0.040$
3C 295 .....	0.456	2	$-0.158 \pm 0.053$
CL 1601+42 .....	0.539	8	$-0.321 \pm 0.043$
CL 0016+16 .....	0.546	7	$-0.281 \pm 0.032$
MS 2053-04 .....	0.583	8	$-0.287 \pm 0.056$
MS 1054-03 .....	0.831	12	$-0.427 \pm 0.040$
RX J0152-13 .....	0.837	13	$-0.449 \pm 0.048$
RX J1226+33 .....	0.892	5	$-0.558 \pm 0.056$
RDCS 1252-29 .....	1.237	4	$-0.586 \pm 0.116$
RDCS 0848+44 .....	1.276	2	$-0.542 \pm 0.069$
SDSS, $n_{\text{den}} < 10$ .....	0.114	1519	$-0.010 \pm 0.031$
Field .....	0.380	8	$-0.164 \pm 0.058$
	0.468	10	$-0.258 \pm 0.061$
	0.563	10	$-0.312 \pm 0.071$
	0.747	10	$-0.445 \pm 0.064$
	0.563	10	$-0.312 \pm 0.071$
	0.747	10	$-0.445 \pm 0.064$
	0.844	10	$-0.473 \pm 0.077$
	0.851	10	$-0.538 \pm 0.053$
	0.951	10	$-0.476 \pm 0.105$
	1.016	10	$-0.626 \pm 0.052$
	1.110	9	$-0.598 \pm 0.055$

slopes of  $-0.454 \pm 0.052$  and  $-0.463 \pm 0.056$ , respectively, consistent with their results. We also note that our treatment of systematic errors is different from previous studies, which leads to slightly different fits (with different quoted uncertainties), even for the same data.

The quoted uncertainty in the fit does not include all sources of systematic error. We first investigate the effects of the choice of coefficients of the FP (which have some uncertainty; see, e.g., Bernardi et al. 2003b; Cappellari et al. 2006). Varying  $a$  from 1.0 to 1.5 produces best-fitting slopes ranging from  $-0.564$  to  $-0.542$ , and varying  $b$  from  $-1.1$  to  $-0.6$  gives slopes in the range  $-0.539$  to  $-0.548$ . We conclude that the precise FP coefficients do not have a large effect on the derived redshift evolution.

Another source of uncertainty is the selection of the clusters. Some are selected in the optical (e.g., CL 1601+42 and CL 0016+16), others in X-rays (e.g., MS 2053-04 and MS 1043-03). Furthermore, the low-redshift Coma Cluster and Bernardi et al. (2003b) sample may not be representative for the descendants of the (very massive) high-redshift clusters in the sample. The data are insufficient to investigate the evolution of subsamples selected by X-ray luminosity, mass, or other parameters. However, we can assess the effects of removing data points and determine whether the fit is driven by some individual clusters. Removing the Bernardi sample changes the best-fitting slope by  $+0.009$  to  $-0.546$ . Removing Coma changes the slope by  $-0.014$ . Removing both Coma and the Bernardi sample changes the slope by  $-0.005$ . The latter test effectively transfers the low-redshift comparison point to the very rich  $z \approx 0.2$  clusters Abell 665, Abell 2218, and Abell 2390, which may be appropriate for the descendants of massive clusters like MS 1054-03 at  $z = 0.83$ . All of these changes are very small and well within the  $1 \sigma$  formal uncertainty in the fit. The cluster with the largest effective weight in the fit is RDCS

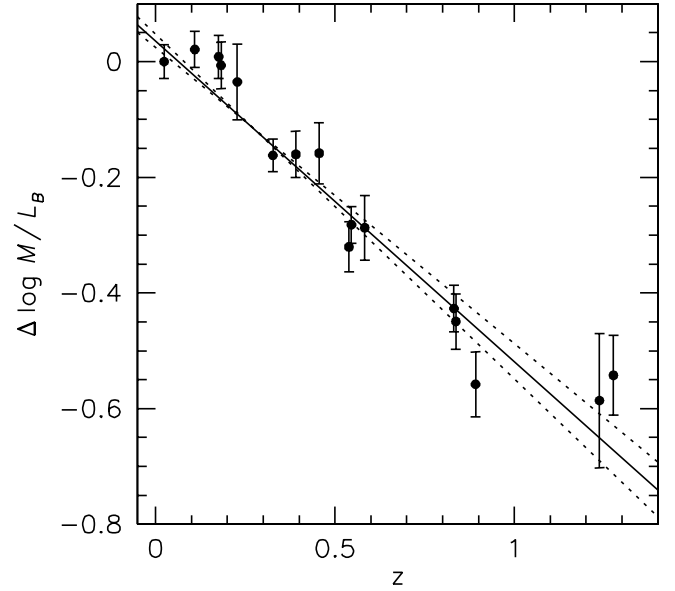


FIG. 6.—Redshift evolution of the mean  $M/L_B$  ratio of cluster galaxies with masses  $M > 10^{11} M_\odot$  in our composite sample. The solid line shows the best-fitting linear function, with slope  $d \log (M/L_B)/dz = -0.555$ . Dotted lines indicate the  $\pm 1 \sigma$  uncertainty in the slope of the relation.

0848+44 at  $z = 1.276$ : removing this cluster changes the slope by  $-0.052$  to  $-0.607$ . The large influence of this single cluster highlights the importance of obtaining more FP measurements for cluster galaxies at  $z > 1$ .

The sample is sufficiently large to determine the scatter in the best-fitting linear relation. The  $\chi^2$  of the fit is 20.4 with 14 degrees of freedom, which implies that a linear function is an adequate description of the data and that the scatter can be explained by the measurement uncertainties. The maximum allowed intrinsic cluster-to-cluster scatter can be determined by requiring that  $\chi^2 > 7.8$ , which for 14 degrees of freedom corresponds to a one-sided probability  $> 0.1$ . Iteratively adding uncertainty due to intrinsic scatter in quadrature to the errors listed in Table 3, we find that  $\chi^2 < 7.8$  for an additional uncertainty of  $> 0.057$ . The 90% confidence upper limit on the intrinsic cluster-to-cluster scatter is therefore 0.057 in  $\log (M/L_B)$ .

## 6. IMPLICATIONS

### 6.1. The Star Formation Epoch of Massive Cluster Galaxies

The observed evolution of the  $M/L_B$  ratio of early-type galaxies depends on cosmological parameters, the star formation histories of the galaxies, their IMF and metallicity, possible changes in dust content with redshift, and selection effects. Although the dependence on cosmological parameters can, in principle, be exploited (see, e.g., Pahre et al. 1996; van Dokkum et al. 1998a; Jørgensen et al. 1999; Lubin & Sandage 2001), we assume here that the evolution of galaxies and their stellar populations is currently more uncertain than the values of  $\Omega_m$  and  $\Omega_\Lambda$  (which are the relevant parameters). We also assume that dust is either absent in early-type cluster galaxies or nonevolving over the redshift range  $0 < z < 1.3$ . The star formation history is parameterized with a single age for the entire stellar population. This is almost certainly incorrect as most early-type galaxies likely have extended and complex star formation histories at early times (e.g., Nagamine et al. 2005; Knudsen et al. 2005; Papovich et al. 2006). However, as shown by, e.g., van Dokkum et al. (1998b), the evolution of a complex stellar population can be well approximated

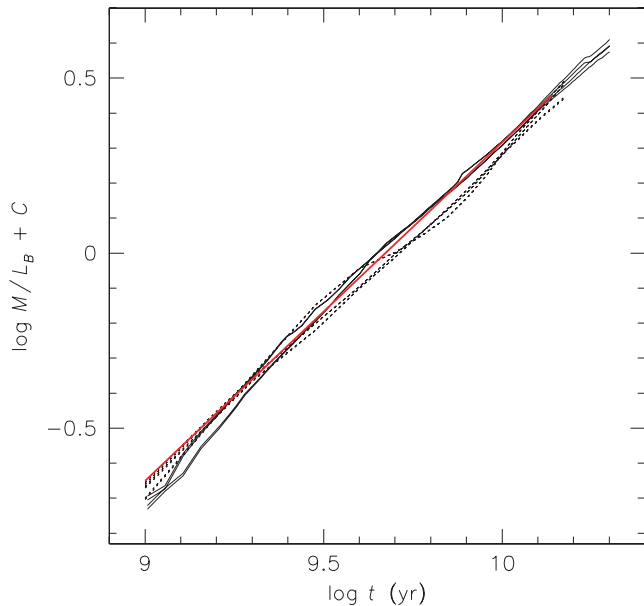


FIG. 7.—Evolution of the  $M/L_B$  ratio of a single-age stellar population, with arbitrary scaling. Solid lines are Bruzual & Charlot (2003) models with different metallicities and IMFs. Dashed lines are Maraston (2005) models with different metallicities, IMFs, and implementation of horizontal branch stars. The predicted evolution is very similar in all models. The red line is the best-fitting power law to a solar metallicity, Salpeter (1955) IMF Bruzual & Charlot (2003) model.

by that of a single-age stellar population of the same luminosity-weighted age, provided that star formation has ceased  $\geq 1$  Gyr before the epoch of observation.

Figure 7 shows the predicted age dependence of the  $M/L_B$  ratio for different assumptions about the IMF, the metallicity, and the implementation of horizontal branch stars in the population synthesis code. Solid lines are various Bruzual & Charlot (2003) models, using the Salpeter (1955) and Chabrier (2003) IMFs and  $[\text{Fe}/\text{H}]$  ranging from solar to +0.56. Dashed lines are various models of Maraston (2005), whose implementation of the late stages of stellar evolution differs from that of Bruzual & Charlot (2003). The Maraston models that are shown are for Salpeter (1955) and Kroupa (2001) IMFs, metallicities ranging from solar to +0.67, and four different implementations of the horizontal branch. All models were normalized in order to emphasize the predicted rate of evolution rather than the absolute  $M/L$  ratios at any given time (which vary greatly between the models).

The predictions are very similar, demonstrating that the predicted evolution is not very dependent on the metallicity, IMF, or particulars of the stellar population synthesis code. An important caveat is that the Salpeter (1955), Kroupa (2001), and Chabrier (2003) IMFs, while different, were all derived from Galactic data, and the IMF of the (progenitors of) early-type galaxies may have been weighted more toward very massive stars (see, e.g., Larson 1998). We return to this issue in § 7.

As first shown by Tinsley (1980), the predicted evolution of the  $M/L$  ratio of a single-age stellar population can be approximated by a power law of the form

$$M/L \propto (t - t_*)^\kappa, \quad (11)$$

with  $t_*$  the formation time of the stars. It can be deduced from the data shown in Figure 7 that for any individual model this approximation is accurate to a few percent over the age range  $9 < \log t < 10$ . Fitting power laws to all of these models gives values of  $\kappa$  in the range 0.93–1.01. A solar metallicity Bruzual &

Charlot (2003) model with a Salpeter (1955) IMF gives a value of 0.97, and in the following we take  $\kappa = 0.97 \pm 0.04$ . We note that the synthetic  $B$  filter adopted by Bruzual & Charlot (2003) (the Buser & Kurucz [1978]  $B_2$  filter) is very similar to the Bessell (1990)  $BX$  filter that we use.

As emphasized by Franx (1993) and van Dokkum & Franx (2001), the observed evolution of the  $M/L_B$  ratio may underestimate the true evolution because of selection effects. If galaxies undergo morphological evolution and transform from late-type galaxies into early-type galaxies at moderate redshift (see, e.g., Dressler et al. 1997), the youngest progenitors of today’s early-type galaxies drop out of the sample at high redshift. This “progenitor bias” leads to biased age estimates, as we trace only the oldest galaxies and not the full population of all progenitor galaxies. The significance of the effect can be estimated from the observed evolution of the early-type galaxy fraction in clusters and from the observed scatter in the color-magnitude relation and FP (see van Dokkum & Franx 2001). The maximum effect on the evolution of the  $M/L_B$  ratio occurs if late-type galaxies are continuously transformed into early-type galaxies and the scatter in the color-magnitude relation is entirely due to age variations. In that extreme case the observed luminosity evolution underestimates the true evolution by  $\Delta_{\text{pbias}} \log(M/L_B) \sim 0.1z$ . The true effect is probably smaller than the maximum value, especially given recent evidence that the evolution of the early-type galaxy fraction out to  $z \sim 1$  is weakest for the most massive objects (Holden et al. 2006). We conservatively assume  $\Delta_{\text{pbias}} \log(M/L_B) = (0.05 \pm 0.05)z$ , which encompasses the full range of possibilities.

With these assumptions and approximations the only free parameter is the time of formation of the stars  $t_*$  (or the corresponding redshift  $z_*$ ). We determined the most likely value and its associated uncertainty using Monte Carlo simulations. In each simulation, each data point listed in Table 3 was perturbed by a value drawn from a Gaussian distribution with a dispersion equal to the uncertainty. Next, progenitor bias was taken into account by decreasing all of the measured  $\log(M/L_B)$  ratios by  $pz$ , with  $p$  drawn from a top-hat probability distribution bounded by 0 and 0.1. Finally, equation (11) was fitted to the perturbed data to obtain  $z_*$ , with the value of  $\kappa$  drawn from a top-hat distribution bounded by 0.93 and 1.01. From 1000 simulations we find that the mean luminosity-weighted formation redshift of the stars in massive cluster early-type galaxies is  $z_* = 2.01^{+0.22}_{-0.17}$ , where the uncertainties indicate the 68% confidence interval. The best-fitting model is shown in Figure 8. Ignoring progenitor bias (i.e., setting  $p = 0$ ) gives  $z_* = 2.23^{+0.24}_{-0.18}$ .

## 6.2. Comparison to Field Galaxies

### 6.2.1. Data

Several recent studies of the fundamental plane of field early-type galaxies have found that the most massive field galaxies out to  $z \sim 1$  have similar  $M/L$  ratios as the most massive cluster galaxies (e.g., Treu et al. 2005a; van der Wel et al. 2005). However, di Serego Alighieri et al. (2006) find that massive cluster galaxies are much older than massive field galaxies, based on a comparison of data from Treu et al. (2005a) (for field galaxies) and Jørgensen et al. (2006) (for cluster galaxies) at  $z \sim 0.9$ . Here we quantify the difference between field and cluster early-type galaxies with  $M > 10^{11} M_\odot$  using our sample of 14 distant clusters and self-consistent modeling of the field and cluster data.

The field samples of van Dokkum et al. (2001), van Dokkum & Ellis (2003), Treu et al. (2005b), and van der Wel et al. (2005) are used. Details are given in Appendix B. Masses and offsets in  $M/L_B$  ratio were calculated in the same way as for the cluster

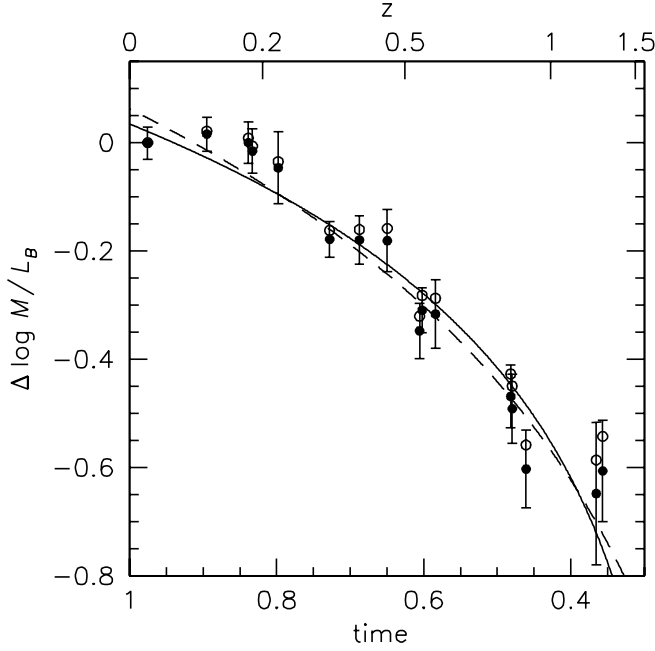


FIG. 8.—Evolution of the mean  $M/L_B$  ratio of massive cluster galaxies with time. Open symbols are the same data points as shown in Fig. 6. Filled symbols with error bars are offset by  $-0.05z$  to account for progenitor bias (see text). The solid line shows the best-fitting model for a Salpeter-like IMF, which has a formation redshift of the stars  $z_* = 2.01$ . The dashed line shows a model with a top-heavy IMF (slope  $x = 0$ ) and a formation redshift  $z_* = 4.0$  (see § 7).

samples (see § 5.1). The sample comprises 87 galaxies with  $M > 10^{11} M_\odot$ , ranging in redshift from 0.32 to 1.14. The sample was divided in bins of 10 galaxies (with the lowest and highest redshift bins containing eight and nine galaxies, respectively), and the central redshift and  $M/L_B$  ratio of each bin were determined using the biweight estimator. The data points are listed in Table 3 and shown in Figure 9, along with the cluster data (not corrected for progenitor bias). The low-redshift point is determined from the Bernardi et al. (2003b) sample (see Appendix A). The  $M/L$  ratios of massive field and cluster galaxies are very similar.

#### 6.2.2. Modeling Approach

The  $M/L$  evolution of field and cluster galaxies can be described by

$$(M/L)_{\text{clus}} = A_c(t - t_c)^{\kappa_c}, \quad (12)$$

$$(M/L)_{\text{field}} = A_f(t - t_f)^{\kappa_f}. \quad (13)$$

The standard approach for determining the star formation epoch of field galaxies is the same as for cluster galaxies: the rate of evolution of the  $M/L$  ratio is used to determine the luminosity-weighted star formation epoch  $t_f$ , and  $A_f$  is a free parameter in the fit (see, e.g., Treu et al. 2002, 2005a, 2005b; van der Wel et al. 2005; di Serego Alighieri et al. 2005). However, this approach is not self-consistent when the evolution of field galaxies is compared to that of cluster galaxies. If field galaxies are younger than cluster galaxies, their mean  $M/L$  ratio evolves faster and is offset from that of cluster galaxies. A self-consistent model that describes the evolution of field and cluster galaxies has  $A_f \equiv A_c$  and has three rather than four free parameters ( $t_c$ ,  $A_c$ , and  $t_f$ ).

The contrast between self-consistent models and standard models is illustrated in Figure 10. The standard approach (*blue dashed lines*) allows models that have different formation redshifts for

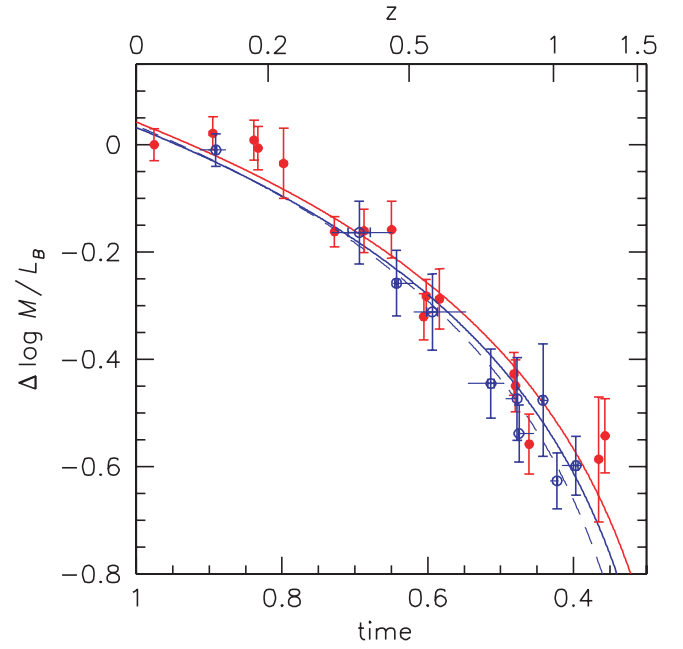


FIG. 9.—Evolution of the mean  $M/L_B$  ratio of field galaxies with  $M > 10^{11} M_\odot$  from the literature (*blue open symbols*) compared to that of cluster galaxies with  $M > 10^{11} M_\odot$  (*red filled symbols*). Thin horizontal lines show the range of redshifts covered in each bin. No corrections for progenitor bias were applied. The red line shows the best-fitting model to the cluster galaxies. The blue dashed line shows a fit to the field galaxies only, and the blue solid line shows a self-consistent fit to the field galaxies given a model for the cluster galaxies (see § 6.2). The age difference between massive field and cluster galaxies is small at  $\sim 4\%$ .

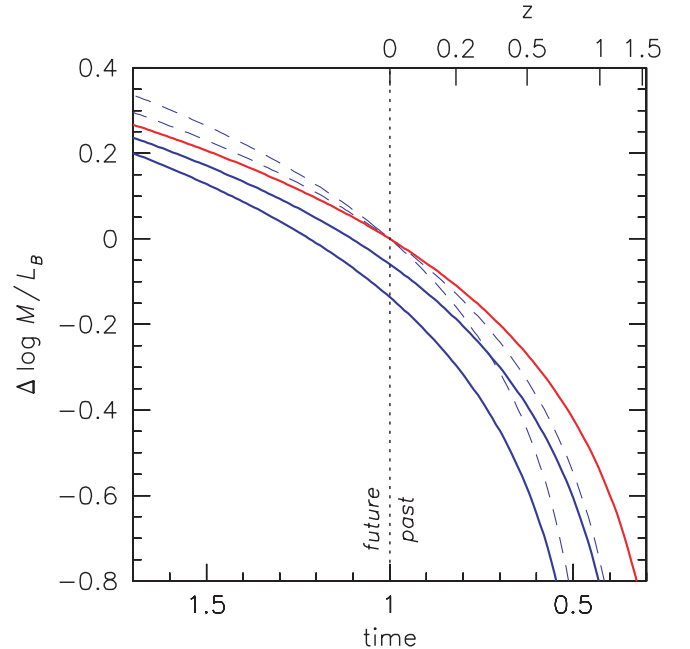


FIG. 10.—Illustration of different ways to model the evolution of the  $M/L_B$  ratio of field galaxies. The red line shows the evolution of a stellar population formed at  $z_* = 2.23$ , appropriate for cluster galaxies. Blue lines show the evolution of field galaxies with  $z_* = 1.5$  (*top*) and  $z_* = 1$  (*bottom*), respectively. Blue dashed lines illustrate the standard fitting approach, which allows models to intersect at  $z = 0$ . Blue solid lines show the self-consistent modeling approach adopted in the present study.

field and cluster galaxies to go through the same point at some redshift, in practice  $z \sim 0$  (see, e.g., Treu et al. 2005b). The implication is that the difference in  $M/L$  ratio between field and cluster galaxies changes sign at  $z \sim 0$ . This is not impossible, as it may be that field galaxies have a different dust content, IMF, metallicity, or progenitor bias from cluster galaxies. However, this would mean that we live at a very special time, namely, the only time in the past or future of the universe when one or more of these effects exactly cancel the difference in  $M/L_B$  ratio resulting from the age difference between field and cluster galaxies. It would also imply that  $\kappa_f \neq \kappa_c$ , as differences in  $A$  would in most models be accompanied by differences in the rate of evolution. The approach followed in this paper is illustrated by the blue solid lines in Figure 10: differences in age result in a different rate of evolution *and* an offset in  $M/L$  ratio, such that the models converge at  $t = \infty$  rather than intersect at the present age of the universe.

### 6.2.3. Results

We determine the star formation epoch of field galaxies  $t_f$  given a model for cluster galaxies ( $t_c$  and  $A_c \equiv A_f$ ). For simplicity we ignore progenitor bias: the relative age of cluster and field galaxies is insensitive to this bias as long as it is similar for both populations. Taking  $z_c = 2.23$  for cluster galaxies and fixing the zero point of the model, we find  $z_f = 1.95^{+0.10}_{-0.08}$  for field galaxies, or an age difference of 4% at the present day. Note that the quoted error does not include the uncertainty in the fit to cluster galaxies. Changing the amount of progenitor bias leads to negligible changes in this result, as long as the bias is the same for field and cluster galaxies.

The significance of this difference is assessed using Monte Carlo simulations, which take the uncertainty in the fit to the cluster galaxies into account. In each simulation the  $M/L_B$  ratios of cluster and field galaxies were perturbed by values drawn from Gaussian distributions with dispersions equal to the uncertainties. Next the formation redshift of cluster galaxies was determined, with the zero point of the model a free parameter in the fit. The formation redshift of field galaxies was also determined, with the zero point of the model fixed to the value that was derived from the cluster galaxies. Finally, the age difference was calculated from the formation redshifts. From 1000 simulations we find that the age difference between massive field and cluster galaxies is  $4.1\% \pm 2.0\%$ . Cluster galaxies are older than field galaxies in 98% of the simulations.

For completeness, we also determined the star formation epoch of field galaxies using the standard approach, i.e., with  $A_f$  as a free parameter. As expected from the lines in Figure 10, the best-fit formation redshift is lower:  $z_f = 1.85^{+0.18}_{-0.13}$ . This model is shown by the blue dashed line in Figure 9.

## 7. DISCUSSION AND CONCLUSIONS

In this paper we presented new measurements of kinematics and structural parameters in three distant galaxy clusters. In vdMvD06a and vdMvD06b the spatially resolved photometric and kinematic profiles of the galaxies are used to measure the evolution of the  $M/L$  ratio independent of many assumptions that enter the FP analysis. In the present study we focused on the FP and the implications for the star formation epoch of the most massive galaxies in clusters and in the field.

We find that the stars in massive early-type galaxies in clusters have a mean luminosity-weighted formation redshift  $z_* = 2.01^{+0.22}_{-0.17}$ . The implied ages are somewhat younger than most previous studies, which were based on smaller samples (e.g., Holden et al. 2005). The quoted error reflects a combination of uncertainties in the data points, the stellar population synthesis models, and the significance of progenitor bias. In particular, varying the amount

of progenitor bias from zero to the maximum allowed by the van Dokkum & Franx (2001) models [ $\sim 0.1z$  in  $\Delta \log (M/L_B)$ ] leads to a range in  $z_*$  of 2.23–1.84. A special form of progenitor bias is “dry” merging, i.e., (nearly) dissipationless mergers of early-type galaxies (e.g., Tran et al. 2005b; van Dokkum 2005; Bell et al. 2006; De Lucia et al. 2006). Numerical simulations suggest that these mergers preserve the edge-on projection of the FP relation (e.g., González-García & van Albada 2003; Boylan-Kolchin et al. 2006; Robertson et al. 2006), which would imply that our analysis is insensitive to this effect.

The main source of uncertainty that our analysis does not (fully) take into account is the form of the IMF. The rate of evolution, as parameterized by  $\kappa$  in our analysis, is very dependent on the logarithmic slope of the IMF in the range  $1\text{--}2 M_\odot$ . Tinsley (1980) finds  $\Delta\kappa \sim -0.3\Delta x$ , where  $x$  is the slope of the IMF, and the more recent Worthey (1994) models give  $\Delta\kappa \sim -0.22\Delta x$ . Although we explored the values of  $\kappa$  for most commonly used IMFs (Salpeter 1955; Kroupa 2001; Chabrier 2003), these are all very similar in the region around  $1 M_\odot$ . These IMFs are probably appropriate for star formation in present-day disk galaxies but may not be applicable to (the progenitors of) early-type galaxies. As is well known, the large mass of heavy elements observed in the hot gas of massive clusters, the abundance ratios of early-type galaxies, numerical simulations of star formation at very high redshift, and other lines of evidence suggest that the IMF may have been top heavy at early times (see, e.g., Worthey et al. 1992; Larson 1998; Schneider et al. 2002 and references therein). The top-heavy IMF preferred by Nagashima et al. (2005) has  $x = 0$  (where  $x = 1.35$  is the Salpeter [1955] slope), which implies  $\kappa_B = 1.4$ . This value of  $\kappa$  gives a best-fitting formation redshift of  $z_* = 4.0$ , higher than the  $z \sim 2$  range derived for Salpeter-like IMFs. This model is shown by the dashed line in Figure 8. We note that our observations probe a very different range of stellar masses than the abundance studies ( $1\text{--}2 M_\odot$  vs.  $\sim 10 M_\odot$ ) and that a top-heavy IMF does not necessarily imply an overall change in the slope but could instead be caused by a time-varying characteristic mass (Larson 1998).

Interestingly, there is some indirect evidence for a top-heavy IMF from the combination of our data with recent results for galaxies at higher redshifts. Although many massive galaxies at  $z \sim 2.5$  are still in the process of forming stars (see § 1), there is also a population with very low specific star formation rates and strong Balmer/4000 Å breaks (e.g., Kriek et al. 2006a, 2006b). These objects formed most of their stars at redshifts  $z \sim 3$  or beyond, inconsistent with the star formation epoch that we derive for Salpeter-like IMFs but in agreement with our results for a top-heavy IMF. Alternative explanations are also possible; e.g., the apparently old  $z \sim 2.5$  galaxies may be “rejuvenated” at later times (see, e.g., Treu et al. 2005b), or they could simply be the oldest objects in a wide distribution of progenitor galaxies.

The age difference between massive field and cluster galaxies is better constrained than their absolute ages, as it is less affected by the uncertainty in the IMF. Our self-consistent modeling approach gives different constraints from the standard approach, as it has fewer free parameters. Assuming that progenitor bias is similar for massive cluster and field galaxies, they have a Salpeter-like IMF, and they have the same metallicity and dust content at a given mass, we find that massive field galaxies are younger than massive cluster galaxies, but that the difference is small at  $4.1\% \pm 2.0\%$  (or  $0.4 \pm 0.2$  Gyr). If a top-heavy IMF with  $x = 0$  is assumed for both field and cluster galaxies, the age difference is  $3.0\% \pm 1.5\%$ .

We note that the quoted errors do not include systematic errors in the comparison of field and cluster galaxies, which are difficult

to quantify. Possible sources of such errors are differences in dynamical structure that deviate from pure homology, selection effects induced by the magnitude limits of the field samples (although these effects are more relevant for low-mass galaxies; see Treu et al. 2005b; van der Wel et al. 2005), a difference in  $\kappa$  and/or  $A$  (which would reflect differences in the IMF or the metallicity), and a difference in the amount of progenitor bias. We tested the effect of a difference in progenitor bias by assuming that cluster early-type galaxies are not affected by it whereas the  $M/L$  ratios of field early-type galaxies need to be corrected by  $-0.1z$ . In that case the age difference increases by a factor of 2, to  $\approx 9\%$ . The importance of progenitor bias is bounded by the observed scatter in the color-magnitude relation of early-type galaxies (van Dokkum & Franx 2001). In clusters this scatter remains low all the way to  $z \sim 1.3$  (Mei et al. 2006), but it is not yet clear whether this also applies to massive field early-type galaxies (see, e.g., Ferreras et al. 2005).

The small age difference we find is consistent with most, but not all, previous studies. It is consistent with the high-redshift field FP studies of Treu et al. (2005b) and van der Wel et al. (2005), the analysis of a sample of gravitational lenses by Rusin & Kochanek (2005), and studies of the local  $Mg_2$ - $\sigma$  relation and FP by Bernardi et al. (1998, 2003b, 2006). However, other studies find larger age differences (always in the sense that field galaxies are younger than cluster galaxies). For nearby massive early-type galaxies, Thomas et al. (2005) and Clemens et al. (2006) find age differences of 1.5–2 Gyr from complex models that include age, metallicity,  $\alpha$ -enhancement, and (in the case of Clemens et al. 2006) carbon enhancement as free parameters. Such large age differences are inconsistent with our analysis at  $>3\sigma$ . Perhaps most importantly, our results are not in agreement with the recent study of di Serego Alighieri et al. (2006), who compare the  $M/L$  ratios of field galaxies at  $z \sim 1$  from Treu et al. (2005b) and di Serego Alighieri et al. (2005) to those of galaxies in two clusters with mean redshift  $\langle z \rangle = 0.86$  from Jørgensen et al. (2006). For galaxies with masses  $>10^{11} M_\odot$ , di Serego Alighieri et al. (2006) find a very large age difference of 3.5–4 Gyr (their Fig. 4). The source of the discrepancy between their results and ours is, at least in part, due to the cluster data that are used in the comparison to field galaxies. As discussed in Appendix A, the  $M/L$  ratios of the galaxies in the Jørgensen et al. (2006) study need to be corrected downward by a factor of  $1 + z$ . Applying this offset to the  $\langle z \rangle = 0.86$  cluster galaxies in Figures 1 and 2 of di Serego Alighieri et al. (2006) brings them in line with the  $z \sim 1$  field galaxies.<sup>6</sup>

The small age difference between massive field and cluster early-type galaxies that we find is remarkable in the context of “standard” hierarchical models, which had predicted that early-

type galaxies in clusters have much older stellar populations than those in the general field (e.g., Baugh et al. 1996; Kauffmann 1996; Kauffmann & Charlot 1998). Specifically, as shown in van Dokkum et al. (2001), the models of Kauffmann et al. (1999) and Diaferio et al. (2001) predicted a systematic offset between field and cluster galaxies of  $\sim 0.25$  in  $\log(M/L_B)$  at all redshifts, clearly at odds with observations. The reason for this environmental dependence in these models is that gas cooling and subsequent star formation continue as long as a galaxy is the central galaxy in its halo. Star formation terminates when a galaxy becomes a satellite in a larger halo, which naturally leads to a large population of “red and dead” galaxies in clusters and blue star-forming galaxies in the field. This inability of standard models to produce red field galaxies is well known, and among proposed solutions is to prevent cooling by postulating that the gas is heated by a central AGN (e.g., Bower et al. 2006; Kang et al. 2006; Croton et al. 2006). The recent study of De Lucia et al. (2006) incorporates such AGN feedback in a very large cosmological simulation. In their simulation, the median star formation epoch  $z_* \sim 2.6$  for elliptical galaxies with masses  $M > 10^{11} M_\odot$ , in good agreement with our results. Furthermore, judging from their Figure 7, the difference in age between elliptical galaxies in halos of  $10^{13}$  and  $10^{15} M_\odot$  is only  $\sim 0.7$  Gyr, consistent with our study. We note that this success does not necessarily imply that the models can reproduce the observed evolution of the FP. Building on the semianalytical models of Baugh et al. (2005) and Bower et al. (2006), Almeida et al. (2006) find that evolution in the radii of early-type galaxies compensates for the evolution in their  $M/L$  ratios, such that the zero point of the FP is approximately constant with redshift, inconsistent with the observations.

Taken at face value, the age difference between field and cluster galaxies is  $\sim 0.4 \pm 0.2$  Gyr, which should be fairly easily detectable at  $z \gtrsim 2$ . Interestingly, this number agrees reasonably well with the age difference that Steidel et al. (2005) find comparing Lyman break galaxies in the general field to those in a protocluster at  $z = 2.3$ . It is also qualitatively consistent with the finding that red galaxies at  $z \sim 2.5$  cluster more strongly than blue galaxies (Quadri et al. 2007).

We thank Arjen van der Wel and Brad Holden for their sleuthing and the referee, Inger Jørgensen, for her comments and generous sharing of unpublished data. Support from NASA grants HF 01126.01-99A, HST AR-09541.01-A, and LTSA NNG04GE12G and from National Science Foundation grant NSF CAREER AST 04-49678 is gratefully acknowledged. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

## APPENDIX A

### LITERATURE DATA FOR CLUSTER GALAXIES

In addition to the three clusters discussed in this paper, we used 11 additional distant clusters in the analysis. We also included the Coma Cluster and a sample of SDSS galaxies in the nearby universe. Here we describe the various transformations that were applied to bring the published data to our system. We also describe sources of systematic uncertainty for each cluster. In most cases morphological information is available; only early-type galaxies with masses  $>10^{11} M_\odot$  were included in the calculation of the  $M/L$  offsets.

#### A1. COMA (ABELL 1656)

The value of  $c_0$  in equation (8) is set to that of the nearby Coma Cluster. This choice is arbitrary, as adding a constant to all values of  $\Delta \log(M/L_B)$  does not change our results in any way. We use the sample of Jørgensen et al. (1996). Structural parameters measured in

<sup>6</sup> See also the erratum published by di Serego Alighieri et al. (2006) in response to our discovery of this discrepancy. Their revised conclusions are consistent with those presented here.



the  $B$  band are given in Jørgensen et al. (1995a), and velocity dispersions corrected to a  $1.7''$  radius aperture are listed in Jørgensen et al. (1995b). The listed surface brightnesses are corrected from the average brightness within the effective radius  $\langle\mu\rangle_e$  to that at the effective radius:  $\mu_e = \langle\mu\rangle_e + 1.393$ . Effective radii are converted from arcseconds to kiloparsecs by adopting a Hubble flow velocity  $v_{\text{flow}} = 7376 \pm 223 \text{ km s}^{-1}$  (see vdMvD06b).

Restricting the sample to early-type galaxies with  $M > 10^{11} M_\odot$ , we find  $c_0 = -9.626$ . The value of  $\Delta \log(M/L_B)$  is zero by definition, with a random uncertainty of  $\pm 0.014$ . The systematic uncertainty in  $\log(M/L_B)$  is a combination of several factors. Zero-point uncertainties in the  $B$ -band photometry give  $\pm 0.010$  (Jørgensen et al. 1992), and the uncertainty in  $v_{\text{flow}}$  implies  $\pm 0.013$ . A comparison of studies by different authors gives  $\pm 0.02$  (see Hudson et al. 2001; vdMvD06b); this uncertainty includes (and may be dominated by) systematic differences in the methodology for deriving structural parameters and velocity dispersions. As discussed in vdMvD06b, several other systematic uncertainties cancel in the comparison to the distant clusters. For example, the uncertainty in the Hubble constant cancels because we are only concerned with the evolution of the  $M/L_B$  ratio and not with the absolute value. Also, as discussed in van Dokkum & Franx (1996), our methodology for measuring velocity dispersions of galaxies in distant clusters mimics that of Jørgensen et al. (1992) in the nearby universe. The combined systematic error is 0.026, and assuming that the random and systematic errors can be added in quadrature, we obtain  $\Delta \log(M/L_B) = 0.000 \pm 0.029$ .

## A2. THE BERNARDI SAMPLE

Bernardi et al. (2003b) have studied the FP relation of galaxies in the SDSS. This sample has several advantages over previous studies: it is large (the total sample comprises 8661 objects), does not suffer from uncertainties due to peculiar velocities (both because they average out over the large volume probed by the data and because of the relatively high typical redshift of 0.1), has very homogeneous spectroscopy and multiband photometry, probes a large range of environments, and spans a range in redshift so that in principle evolutionary effects can be studied within the sample. An important drawback specific to this sample is that the magnitude selection causes significant biases that need to be taken into account (Bernardi et al. 2003b); such biases are less important (although not absent) for cluster samples as all galaxies are at the same distance. Furthermore, as noted by, e.g., Cappellari et al. (2006), the coefficients  $a$  and  $b$  that Bernardi et al. (2003b) derive do not appear to be consistent with most other studies of the FP; the cause for this discrepancy is unclear at present.

Velocity dispersions and structural parameters in the rest-frame  $g'$  band were obtained from Table 3b of Bernardi et al. (2003a). Effective radii were converted to our cosmology. Rest-frame  $g'$  band surface brightnesses were converted to the rest-frame  $B$  band using the transformation

$$B = g' + 0.44(g' - r') + 0.17, \quad (\text{A1})$$

with  $B$  on the Vega system and  $g'$  and  $r'$  on the AB system. For consistency this transformation was derived using the same method and  $B$ -band filter curve as equation (1), with the only difference that the redshift was set to zero (i.e., the  $g'$  and  $r'$  filters were not redshifted). Fukugita et al. (1996) give  $B_z = g' + 0.42(g' - r') + 0.22$ ; for the typical colors of galaxies in the sample this transformation is consistent with ours to 0.04 mag. The  $g' - r'$  colors cannot be obtained by simply taking the difference  $\langle\mu_{g'}\rangle_e - \langle\mu_{r'}\rangle_e$  as the surface brightness was evaluated within a different radius in each band. Therefore, we first corrected the surface brightnesses to a common radius using the measured effective radii of the galaxies (see van der Wel et al. 2005). Finally, we applied the transformation  $\mu_e = \langle\mu\rangle_e + 1.393$ . We note that the uncertainty in this procedure is ultimately determined by the difference between the observed band and the final rest-frame band. In this case, the observed band is the  $g'$  band and the final rest-frame band is the  $B$  band, and at  $z \sim 0.1$  these are very well matched.

Bernardi et al. (2003a) corrected their velocity dispersions to an aperture of radius  $0.125r_e$ , which for almost all galaxies is significantly smaller than the  $1.7''$  aperture at the distance of Coma that we adopted. We first retrieved the dispersions as measured through the  $1.5''$  SDSS fibers using the measured effective radii (in arcseconds) and equation (1) of Bernardi et al. (2003a). Then we corrected these measured dispersions to a  $1.7''$  aperture at the distance of Coma using the redshifts of the galaxies and the prescription of Jørgensen et al. (1995b).

FP offsets were determined for all galaxies in the sample with local density  $n_{\text{den}} > 20$ , using  $a = 1.20$ ,  $b = -0.83$ , and  $c_0 = -9.626$ . Following Bernardi et al. (2003a) (their Fig. 10) and Bernardi et al. (2003b) (their Fig. 9), galaxies with  $n_{\text{den}} = 100$  were excluded. As discussed extensively in Bernardi et al. (2003b), average  $M/L$  ratios inferred from these data need to be corrected for selection effects. Fitting to the values listed in their Figure 7 gives  $\Delta_{\text{select}} \log(M/L_{g'}) = -0.289z$  and  $\Delta_{\text{select}} \log(M/L_{r'}) = -0.208z$ . Extrapolating to the  $B$  band using equation (A1) gives  $\Delta_{\text{select}} \log(M/L_B) = -0.320z$ . Bernardi et al. (2003b) determined the effect for the full sample, and not for the subsample of galaxies with  $M > 10^{11} M_\odot$  that is relevant here. Selection effects will be less important for the massive galaxies, but it is not clear by how much. We assume that the effect is half that of the full sample and conservatively assign an uncertainty that encompasses the full range of possibilities:

$$\Delta \log(M/L_{B,\text{corr}}) = \Delta \log(M/L_B) + (0.16 \pm 0.16)z. \quad (\text{A2})$$

Given the uncertainty in the redshift evolution within the Bernardi et al. (2003b) sample, we evaluate the average  $M/L_B$  offset at one redshift only. We chose  $z \approx 0.10$ , as selection effects do not play a large role at that redshift and the local galaxy densities  $n_{\text{den}}$  are not well determined for  $z \lesssim 0.08$  (Bernardi et al. 2003a). In practice, we select the 171 galaxies with  $M > 10^{11} M_\odot$ ,  $20 < n_{\text{den}} < 100$ , and  $0.075 < z < 0.14$ . The average redshift of this sample is  $z = 0.109$ .

The value of the  $M/L$  offset is listed in Table 3. The random uncertainty is only  $\pm 0.008$ . The uncertainty in the photometric transformation is taken to be the difference between the Fukugita et al. (1996) transformation and ours ( $\pm 0.016$ ). The uncertainty due to

selection effects is  $0.16 \times 0.109 = 0.017$ . Bernardi et al. (2003a) estimate that the systematic uncertainty in the velocity dispersions is  $\sim 3\%$ , implying  $\pm 0.019$  in  $\log(M/L_B)$ . Adding all uncertainties in quadrature gives  $\pm 0.030$  for the total error in the  $M/L$  offset.

For the analysis in § 6.2 we also determined the  $M/L$  offset in low-density regions, selecting all 1519 galaxies with  $M > 10^{11} M_\odot$ ,  $0.075 < z < 0.14$ , and  $n_{\text{den}} < 10$ . The average redshift of this sample is 0.1093. The offset (corrected for selection effects) is listed in Table 3. The random error is 0.003 and the systematic error is the same as derived for the high-density sample. The difference between the high- and low-density offset is 0.031 in  $\log(M/L_B)$ , or 0.08 mag if it is interpreted as a difference in surface brightness at fixed  $r_e$  and  $\sigma$ . This difference appears to be consistent with the trends shown in Figure 9 of Bernardi et al. (2003b) for this redshift range.

### A3. ABELL 2218

Abell 2218 is a well-studied, very rich cluster at  $z = 0.1756$  (Le Borgne et al. 1992). Its FP has been analyzed in two independent studies, Jørgensen et al. (1999) and Ziegler et al. (2001). Jørgensen et al. (1999) give velocity dispersions and two sets of structural parameters, one derived from *HST* images and one derived from ground-based data. Their velocity dispersions have already been corrected to the  $1.7''$  aperture at the distance of Coma. The ground-based structural parameters were corrected from the observed  $I_c$  band to the rest-frame  $B$  band using the listed  $V - I$  colors:

$$B_z = I_c + 1.29(V - I_c) + 0.18, \quad (\text{A3})$$

and corrected from  $\langle \mu \rangle_e$  to  $\mu_e$ . The structural parameters measured from the *HST* F702W images were transformed by Jørgensen et al. (1999) to the  $I_c$  band, and we use equation (A3) to transform the surface brightness to rest-frame  $B$ .

Ziegler et al. (2001) obtained independent spectroscopy for galaxies in Abell 2218 and give velocity dispersions that were corrected to the same  $1.7''$  radius aperture at the distance of Coma as we use. They used the same *HST* imaging as Jørgensen et al. (1999) but derive their own structural parameters from the images. We transformed their listed rest-frame Gunn  $r$  surface brightnesses from  $L_\odot \text{ pc}^{-2}$  to  $\text{mag arcsec}^{-2}$  using their equation (7) and then converted them back to observed  $I_c$  magnitudes using their equation (1). In this conversion we used the listed  $V$  and  $I_c$  magnitudes after applying extinction corrections of 0.083 mag in  $V$  and 0.048 mag in  $I_c$ . The  $I_c$  surface brightnesses were transformed from  $\langle \mu \rangle_e$  to  $\mu_e$  and corrected to rest-frame  $B$  using equation (A3).

Combining all information, there are three measurements of the FP: (1) dispersions and *HST* structural parameters from Jørgensen et al. (1999), (2) dispersions and ground-based structural parameters from Jørgensen et al. (1999), and (3) dispersions and *HST* structural parameters from Ziegler et al. (2001). The two sets of dispersions and the two sets of ground-based data are independent. We determined the  $M/L$  offset for all three cases and compared the results. The offsets agree very well: differences are approximately 0.02 in  $\log(M/L_B)$ . Rather than take the average of the three determinations, we use the zero point from measurement 2: it falls between the other two and is the one used in the analysis of the Jørgensen et al. (1999) paper. Apart from a systematic uncertainty of 0.02 due to the differences between the three determinations, there is also an uncertainty introduced by the fact that all structural parameters were measured in redder bands than the redshifted  $B$  band. This uncertainty stems from color gradients and the dependence of the FP parameters on passband and is estimated at 0.05 mag. The combined systematic uncertainty is 0.029 in  $\log(M/L_B)$ .

### A4. ABELL 665

The FP in Abell 665 ( $z = 0.1829$ ; Gómez et al. 2000) was studied by Jørgensen et al. (1999). As for Abell 2218, there are two sets of structural parameters, one from *HST* imaging and the other from ground-based imaging. Contrary to Abell 2218, there are also two sets of  $V - I_c$  color measurements: one directly measured from the ground and the other transformed from  $R_{606} - I_{814}$  *HST* colors. We applied the same transformations as for Abell 2218 and compared the FP relations derived from ground-based data to the FP derived from *HST* data. The difference is 0.02 in  $\log(M/L_B)$ , consistent with the expected uncertainties due to the color transformations. The final offset was calculated from the *HST* measurements in Jørgensen et al. (1999). As for Abell 2218, the total systematic uncertainty is estimated at 0.029 in  $\log(M/L_B)$ .

### A5. ABELL 2390

Fritz et al. (2005) present the FP relation in Abell 2390 at  $z = 0.228$ . We corrected the listed velocity dispersions to our fiducial aperture following Jørgensen et al. (1995b), using  $1.5'' \times 2.8''$  as the extraction aperture. The listed rest-frame Gunn  $r$  surface brightnesses were converted back to observed WFPC2  $I_{814}$  magnitudes using  $\langle \mu_{814} \rangle_e = \langle \mu_{r,z} \rangle_e - 0.75$  (see Fritz et al. 2005). We converted  $\langle \mu \rangle_e$  to  $\mu_e$  and transformed to the (COSMIC camera)  $I$  band using (Holtzman et al. 1995)

$$I = I_{814} - 0.037(B - I) + 0.007(B - I)^2 + 0.00. \quad (\text{A4})$$

The  $B - I$  colors were determined from the listed  $B$  and  $I$  aperture magnitudes. These are not corrected for extinction; we applied corrections of 0.476 mag to  $B$  and 0.214 mag to  $I$  (see Fritz et al. 2005). Rest-frame  $B$  surface brightnesses were determined using

$$B_z = I + 0.47(B - I) + 0.51. \quad (\text{A5})$$

Based on listed zero-point uncertainties, the uncertainty in the photometric transformations is estimated at 0.03 mag.

In contrast to our own measurements of the FP in distant clusters, no effort was made to determine parameters in the same way as the Jørgensen et al. (1995b) studies at low redshift. There is therefore a systematic uncertainty of  $\approx 5\%$  in the velocity dispersions, which



is caused by possible differences in the fitting region, choice of templates, and fitting methodology. This corresponds to an uncertainty of 0.031 in  $\log(M/L_B)$ . Combined with a 0.05 mag uncertainty due to the transformations and the fact that the FP was determined in red rest-frame bands, we estimate that the total systematic uncertainty is 0.037 in  $\log(M/L_B)$ .

#### A6. CL 1358+62

Kelson et al. (2000a, 2000b, 2000c) provide an extensive analysis of the FP in the cluster CL 1358+62 at  $z = 0.327$  in the rest-frame  $V$  band. Listed velocity dispersions (Kelson et al. 2000b) are already corrected to our fiducial aperture. Listed rest-frame surface brightnesses and colors were back-corrected to the observed ones using the equations in Kelson et al. (2000a). They were then transformed to rest-frame  $B$  using

$$B_z = I_{814} + 1.09(R_{606} - I_{814}) + 0.52 \quad (\text{A6})$$

and transformed from  $\langle\mu\rangle_e$  to  $\mu_e$ .

The Kelson et al. (2000b) methodology is essentially identical to ours, and systematic uncertainties in, e.g., spectral continuum filtering cancel. Systematic uncertainties in the dispersions due to possible changes of the spectral templates with redshift are estimated at 2%. Systematic errors in the photometric transformations are estimated at 0.03 mag. The combined systematic uncertainty is 0.021 in  $\log(M/L_B)$ .

#### A7. CL 0024+16

Van Dokkum & Franx (1996) present the FP in the  $z = 0.391$  cluster CL 0024+16. We use their sample rather than the much larger sample of Moran et al. (2005) because van Dokkum & Franx (1996) use identical techniques to ours, and systematic uncertainties in the velocity dispersions dominate over random uncertainties due to the limited sample size. We note that Moran et al. (2005) find that their data agree with those of van Dokkum & Franx (1996) to within a few percent, both for the velocity dispersions of individual objects in common between the two samples and for the offset in  $M/L$  ratio derived from the FP.

Listed velocity dispersions have already been corrected to our fiducial aperture. The listed  $I$ -band surface brightnesses were corrected to rest-frame  $B$  using

$$B_z = I + 1.46(R - I) + 0.57. \quad (\text{A7})$$

The systematic uncertainty stems from the same sources as for CL 1358+62 and is estimated at 0.021 in  $\log(M/L_B)$ .

#### A8. MS 2053–04 AND MS 1054–03

Wuyts et al. (2004) analyze the FP in the two clusters MS 2053–04 ( $z = 0.583$ ) and MS 1054–03 ( $z = 0.832$ ). The Wuyts et al. (2004) study followed initial studies of smaller samples in the two clusters by Kelson et al. (1997) and van Dokkum et al. (1998a), respectively. The listed structural parameters and velocity dispersions are already on our system. However, we update the transformations to rest-frame  $B$ , as the synthetic  $B$  filter curve used in Wuyts et al. (2004) is slightly different from the one used in the present study. The updated transformations are

$$B_z = I_{814} + 0.38(R_{606} - I_{814}) + 0.93 \quad (\text{A8})$$

for MS 2053–04 and

$$B_z = I_{814} - 0.05(R_{606} - I_{814}) + 1.22 \quad (\text{A9})$$

for MS 1054–03. These transformations are consistent with those given in Wuyts et al. (2004) to within  $\approx 0.04$  mag for the typical colors of early-type galaxies in these clusters. The systematic uncertainty is taken to be the same as for CL 1358+62.

#### A9. RX J0152.7–1357

Imaging and spectroscopic data of RX J0152–13 ( $z = 0.837$ ) and RX J1226+33 ( $z = 0.892$ ) were obtained in the context of the Gemini/*HST* Galaxy Cluster Project (Jørgensen et al. 2005). Jørgensen et al. (2006) discuss the FP in these clusters, finding that the most massive galaxies in these clusters are offset from the Coma FP by  $\Delta\log(M/L_B) \approx -0.25$  only. This offset is much smaller than had been found previously for the cluster MS 1054–03 at  $z = 0.83$  (Wuyts et al. 2004), indicating significant cluster-to-cluster scatter at  $z \sim 1$ . Jørgensen et al. (2006) do not provide data for individual galaxies in these clusters, but for RX J0152–13 the FP can be constructed by combining velocity dispersions listed in Table 12 of Jørgensen et al. (2005) with structural parameters measured by Blakeslee et al. (2006). The velocity dispersions ( $\sigma_{\text{corr}}$  from Jørgensen et al. 2005) are already on our system, and no further corrections are required. Effective radii from Blakeslee et al. (2006) were circularized by multiplying them by  $(b/a)^{1/2}$  and converted to kiloparsecs. The listed magnitudes ( $i_{775, \text{gfit}}$  in Blakeslee et al. 2006) were converted to surface brightnesses on our system through the transformation

$$\mu_{e,775} = i_{775, \text{gfit}} + 2.5 \log 2 + 2.5 \log(\pi r_e^2) + 1.393, \quad (\text{A10})$$

with  $r_e$  in arcseconds. Next, surface brightnesses and colors were converted from AB to Vega magnitudes using the Sirianni et al. (2005) ACS zero points and transformed to rest-frame  $B$  using

$$B_z = i_{775} - 0.35(i_{775} - z_{850}) + 1.21. \quad (\text{A11})$$

Finally, the  $M/L$  offset was derived using the same techniques as for the other clusters. We find  $\Delta \log (M/L_B) = -0.449$ , much larger than the Jørgensen et al. (2006) offset. This difference is caused by a combination of two effects: small differences in methodology (e.g., a 0.05 mag difference between the measured Blakeslee et al. [2006] and Jørgensen et al. [2006]  $i_{775}$  surface brightnesses) and an error in the transformation from observed to rest-frame magnitudes. The rest-frame magnitudes in Jørgensen et al. (2006) are too faint by a factor of  $1 + z$ .<sup>7</sup> I. Jørgensen kindly determined the offset for galaxies with  $M > 10^{11} M_\odot$  in RX J0152–13 using the Jørgensen et al. (2006) data with our transformation from SDSS  $i'$  and  $z'$  magnitudes to rest-frame  $B$  and our zero point for the Coma Cluster. The offset is  $-0.475$ . This value is consistent with ours to  $<0.01$ , when the systematic difference of 0.05 mag between the Blakeslee et al. (2006) and Jørgensen et al. (2006) methodology is taken into account. Sources of systematic error are the same as for, e.g., Abell 2218 and Abell 665, but we added 0.02 in quadrature to reflect the difference between the Blakeslee et al. (2006) and Jørgensen et al. (2006) structural parameters.

#### A10. RX J1226.9+3332

This  $z = 0.892$  cluster is discussed in Jørgensen et al. (2006) together with RX J0152–13. No individual measurements of galaxies are available, but I. Jørgensen determined the offset in  $M/L_B$  ratio using our transformations, Coma zero point, and mass selection. A small correction of 0.05 mag was applied to bring the data to the same system as the Blakeslee et al. (2006) photometry (see § A9). The offset is consistent with the value that is implied by the Jørgensen et al. (2006) study, after correcting their data points by a factor  $1 + z$  (§ A9). Sources of systematic uncertainty are the correction to the Blakeslee et al. (2006) system [0.02 in  $\log (M/L_B)$ ], systematic uncertainty in the dispersions (0.025), the transformation to rest-frame  $B$  (0.012), and the photometric zero points of the ground-based photometry (0.02). The total systematic uncertainty is 0.040. The random uncertainty is 0.039 in  $\log (M/L_B)$ , which brings the total error to 0.056.

#### A11. RDCS 1252.9–2927

The FP in the  $z = 1.237$  cluster RDCS 1252 was determined by Holden et al. (2005) using four galaxies, all of which have  $M > 10^{11} M_\odot$ . Listed velocity dispersions are already on our system. The observed  $z_{850}$  surface brightnesses were converted from  $\langle \mu \rangle_e$  to  $\mu_e$  and converted to rest-frame  $B$  using

$$B_z = z_{850} - 0.22(i_{775} - J) + 0.99, \quad (\text{A12})$$

where  $z_{850}$ ,  $i_{775}$ , and  $J$  are on the AB system (see Holden et al. 2005). This transformation is slightly different from the one used in Holden et al. (2005), due to the use of a different derivation procedure.<sup>8</sup> For the typical colors of galaxies in RDCS 1252 the difference in the transformation results in a difference of  $\lesssim 0.05$  mag in  $B_z$ . The determination of velocity dispersions from the near-UV spectral region may cause systematic errors of 5% (van Dokkum & Stanford 2003). The total systematic uncertainty is 0.043 in  $\log (M/L_B)$ . We note that for this cluster and for RDCS 0848 the random uncertainty exceeds the systematic uncertainty.

#### A12. RDCS 0848+4453

The FP for three galaxies in the  $z = 1.276$  cluster RDCS 0848 was presented in van Dokkum & Stanford (2003). The listed velocity dispersions are on our system. The F160W surface brightnesses were corrected to rest-frame  $B$  using

$$B_z = H_{160} + 0.46(I_{814} - H_{160}) + 1.92. \quad (\text{A13})$$

This transformation is slightly different from equation (3) in van Dokkum & Stanford (2003), as we use a slightly different synthetic  $B$  filter in the present study. For the typical colors of galaxies in RDCS 0848 the two transformations are consistent to  $\lesssim 0.05$  mag. The systematic uncertainty is a combination of the uncertainty derived above for CL 1358+62 and a 5% uncertainty in the dispersions, which originates from the use of the near-UV spectral region. The adopted uncertainty in  $\log (M/L_B)$  is 0.038.

## APPENDIX B

### LITERATURE DATA FOR FIELD GALAXIES

#### B1. VAN DOKKUM ET AL. (2001) AND VAN DOKKUM & ELLIS (2003)

These studies give structural parameters and velocity dispersions for morphologically selected early-type galaxies. Van Dokkum et al. (2001) give results for 18 galaxies at  $0.15 \leq z \leq 0.55$  in the foreground of the rich clusters MS 2053–04 ( $z = 0.58$ ) and MS 1054–03

<sup>7</sup> See the erratum published by Jørgensen et al. (2006) in response to our discovery of this discrepancy.

<sup>8</sup> Note that most of the apparent difference between eq. (A12) and the transformation given in Holden et al. (2005) results from the fact that in their equation the  $z_{850}$  surface brightness is on the AB system and the  $i_{775} - J$  color is on the Vega system.

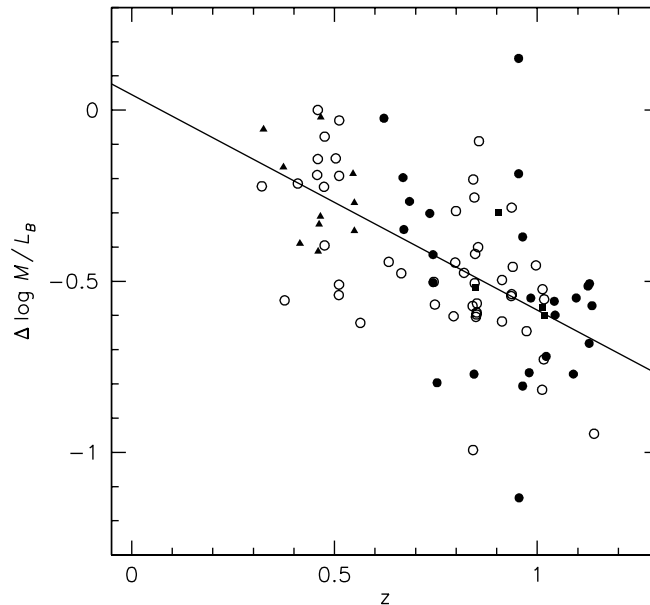


FIG. 11.—Individual field early-type galaxies with  $M > 10^{11} M_{\odot}$  in the four samples that are combined in this study. Filled triangles are galaxies from van Dokkum et al. (2001), filled squares are from van Dokkum & Ellis (2003), filled circles are from van der Wel et al. (2005), and open circles are from Treu et al. (2005b).

( $z = 0.83$ ). Van Dokkum & Ellis (2003) studied 10 galaxies at  $0.56 \leq z \leq 1.10$  in the Hubble Deep Field–North (HDF-N). The methodology for deriving velocity dispersions and structural parameters in these studies is identical to ours. Listed surface brightnesses were transformed to rest-frame  $B$  following the procedures outlined in the cited papers. Note that this transformation is different for each galaxy, as it depends on the redshift. Effective radii were converted to kpc. We determined masses and offsets in  $M/L_B$  ratio in the same way as for individual cluster galaxies (see § 5.1). Eleven of 18 galaxies in the MS 1054–03 field, and 4 of 10 galaxies in the HDF-N, have masses  $> 10^{11} M_{\odot}$ .

#### B2. VAN DER WEL ET AL. (2005)

Van der Wel et al. (2005) study a sample of 38 field early-type galaxies in two fields, the Chandra Deep Field–South (CDF-S) and the field containing the cluster RDCS 1252–29 ( $z = 1.24$ ). The galaxies were selected by a combination of color and morphological criteria. They span the redshift range  $0.62 \leq z \leq 1.14$ , with median 0.97. Van der Wel et al. (2005) follow the same procedures as we do here for determining structural parameters and velocity dispersions. The listed surface brightnesses (Table 2 in van der Wel et al. 2005) were transformed to the rest-frame  $B$  band, and effective radii were converted to kiloparsecs. Twenty-five galaxies in this sample have  $M > 10^{11} M_{\odot}$ .

#### B3. TREU ET AL. (2005a, 2005b)

Treu et al. (2005a, 2005b) study a very large sample of 226 visually classified bulge-dominated galaxies in the GOODS HDF-N region. Although Treu et al. (2005a, 2005b) use different software and procedures from those we use here, a direct comparison between their data and those of van Dokkum & Ellis (2003) (which are on our system) shows that any systematic differences are very small. Specifically, for a subset of the galaxies velocity dispersions were determined using the same software as used in the present study (see § 2.5), and the two sets of measurements agree to  $\lesssim 1\%$  on average (see Treu et al. 2005b). Surface brightnesses were converted to rest-frame  $B$  following the procedure outlined in Treu et al. (2005b). Next, they were converted from  $\langle \mu \rangle_e$  to  $\mu_e$ , and from the AB system to the Vega system. Effective radii were converted to kiloparsecs. Treu et al. (2005b) correct the measured velocity dispersions for aperture effects by applying a fixed correction of  $\sigma_{\text{corr}} = 1.10 \sigma_{\text{ap}}$ . For consistency we correct the measured  $\sigma_{\text{ap}}$  values to the equivalent of a  $3.4''$  diameter aperture at the distance of the Coma Cluster (see § 2.5).

Masses and offsets in  $M/L$  ratio were calculated following the procedures of § 5.1. Of 141 galaxies classified as early-type, 25 were excluded because they could be misclassified spiral galaxies, as judged from the residuals from an  $r^{1/4}$  fit (see § 2.1 in Treu et al. 2005b). Of the remaining 116 galaxies, 48 have  $M > 10^{11} M_{\odot}$ . These galaxies are at  $0.32 \leq z \leq 1.14$ , with median 0.84.

#### B4. COMPARISON OF INDIVIDUAL SAMPLES

The total number of field early-type galaxies with  $M > 10^{11} M_{\odot}$  is 88. One galaxy was excluded from this sample: object 1 from van Dokkum et al. (2001). This galaxy is at  $z = 0.183$ , and as the next lowest redshift is  $z = 0.321$ , including this object would “artificially” create a very wide bin at low redshift in Figure 9.<sup>9</sup> In § 6.2 we treat the remaining 87 galaxies as a single sample; here we test whether this approach is warranted, by examining whether there is evidence for systematic differences between the four samples.

<sup>9</sup> We note that this object follows the trend defined by the other 87 galaxies.

Data points for individual galaxies are shown in Figure 11. We determined a mean offset for each sample in the following way. We fitted a linear function to the binned field data listed in Table 3. This function has the form  $\Delta \log(M/L_B) = 0.045 - 0.629z$ . After subtracting this linear function from the individual data points, the average of the residuals was determined for each of the four literature samples using the biweight estimator. We find the following average offsets in  $\log(M/L_B)$ :  $-0.01 \pm 0.01$  (van Dokkum et al. 2001),  $-0.01 \pm 0.04$  (van Dokkum & Ellis 2003),  $+0.02 \pm 0.04$  (van der Wel et al. 2005), and  $-0.01 \pm 0.02$  (Treu et al. 2005b). These values are consistent with each other, and we conclude that it is appropriate to treat the four samples as a single, large sample.

## REFERENCES

- Almeida, C., Baugh, C. M., & Lacey, C. G. 2006, MNRAS, submitted (astro-ph/0608544)
- Baugh, C. M., Cole, S., & Frenk, C. S. 1996, MNRAS, 283, 1361
- Baugh, C. M., Lacey, C. G., Frenk, C. S., Granato, G. L., Silva, L., Bressan, A., Benson, A. J., & Cole, S. 2005, MNRAS, 356, 1191
- Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
- Bell, E. F., et al. 2006, ApJ, 640, 241
- Bernardi, M., Nichol, R. C., Sheth, R. K., Miller, C. J., & Brinkmann, J. 2006, AJ, 131, 1288
- Bernardi, M., Renzini, A., da Costa, L. N., Wegner, G., Alonso, M. V., Pellegrini, P. S., Rit , C., & Willmer, C. N. A. 1998, ApJ, 508, L143
- Bernardi, M., et al. 2003a, AJ, 125, 1817
- . 2003b, AJ, 125, 1866
- Bessell, M. S. 1990, PASP, 102, 1181
- Blakeslee, J. P., et al. 2006, ApJ, 644, 30
- Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2006, MNRAS, 370, 645
- Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2006, MNRAS, 369, 1081
- Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
- Buser, R., & Kurucz, R. L. 1978, A&A, 70, 555
- Cappellari, M., et al. 2006, MNRAS, 366, 1126
- Chabrier, G. 2003, PASP, 115, 763
- Clemens, M. S., Bressan, A., Nikolic, B., Alexander, P., Annibali, F., & Rampazzo, R. 2006, MNRAS, 370, 702
- Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
- Croton, D. J., et al. 2006, MNRAS, 365, 11
- Daddi, E., et al. 2003, ApJ, 588, 50
- . 2004, ApJ, 600, L127
- De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499
- Diaferio, A., Kauffmann, G., Balogh, M. L., White, S. D. M., Schade, D., & Ellingson, E. 2001, MNRAS, 323, 999
- di Serego Alighieri, S., Lanzoni, B., & J rgensen, I. 2006, ApJ, 647, L99
- di Serego Alighieri, S., et al. 2005, A&A, 442, 125
- Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
- Dressler, A., et al. 1997, ApJ, 490, 577
- Faber, S. M., Dressler, A., Davies, R. L., Burstein, D., Lynden-Bell, D., Terlevich, R., & Wegner, G. 1987, in *Nearly Normal Galaxies*, ed. S. M. Faber (New York: Springer), 175
- Ferreras, I., Lisker, T., Carollo, C. M., Lilly, S. J., & Mobasher, B. 2005, ApJ, 635, 243
- F rster Schreiber, N. M., et al. 2004, ApJ, 616, 40
- Franx, M. 1993, PASP, 105, 1058
- Franx, M., et al. 2003, ApJ, 587, L79
- Freedman, W. L., et al. 2001, ApJ, 553, 47
- Fritz, A., Ziegler, B. L., Bower, R. G., Smail, I., & Davies, R. L. 2005, MNRAS, 358, 233
- Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
- G mez, P. L., Hughes, J. P., & Birkinshaw, M. 2000, ApJ, 540, 726
- Gonz lez-Garc a, A. C., & van Albada, T. S. 2003, MNRAS, 342, L36
- Grazian, A., et al. 2006, A&A, 453, 507
- Guzman, R., Koo, D. C., Faber, S. M., Illingworth, G. D., Takamiya, M., Kron, R. G., & Bershad, M. A. 1996, ApJ, 460, L5
- Holden, B. P., et al. 2005, ApJ, 620, L83
- . 2006, ApJ, 642, L123
- Holtzman, J. A., Burrows, C. J., Casertano, S., Hester, J. J., Trauger, J. T., Watson, A. M., & Worthey, G. 1995, PASP, 107, 1065
- Hudson, M. J., Lucey, J. R., Smith, R. J., Schlegel, D. J., & Davies, R. L. 2001, MNRAS, 327, 265
- J rgensen, I., Bergmann, M., Davies, R., Barr, J., Takamiya, M., & Crampton, D. 2005, AJ, 129, 1249
- J rgensen, I., Chiboucas, K., Flint, K., Bergmann, M., Barr, J., & Davies, R. 2006, ApJ, 639, L9
- J rgensen, I., Franx, M., Hjorth, J., & van Dokkum, P. G. 1999, MNRAS, 308, 833
- J rgensen, I., Franx, M., & Kjaergaard, P. 1992, A&AS, 95, 489
- . 1995a, MNRAS, 273, 1097
- . 1995b, MNRAS, 276, 1341
- . 1996, MNRAS, 280, 167
- Kang, X., Jing, Y. P., & Silk, J. 2006, ApJ, 648, 820
- Kauffmann, G. 1996, MNRAS, 281, 487
- Kauffmann, G., & Charlot, S. 1998, MNRAS, 297, L23
- Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 303, 188
- Kelson, D. D. 2003, PASP, 115, 688
- Kelson, D. D., Illingworth, G. D., van Dokkum, P. G., & Franx, M. 2000a, ApJ, 531, 137
- . 2000b, ApJ, 531, 159
- . 2000c, ApJ, 531, 184
- Kelson, D. D., van Dokkum, P. G., Franx, M., Illingworth, G. D., & Fabricant, D. 1997, ApJ, 478, L13
- Knudsen, K. K., et al. 2005, ApJ, 632, L9
- Kriek, M., et al. 2006a, ApJ, 645, 44
- . 2006b, ApJ, 649, L71
- Krist, J. 1995, in ASP Conf. Ser. 77, *Astronomical Data Analysis Software and Systems IV*, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco: ASP), 349
- Kroupa, P. 2001, MNRAS, 322, 231
- Lab e, I., et al. 2003, AJ, 125, 1107
- . 2005, ApJ, 624, L81
- Larson, R. B. 1998, MNRAS, 301, 569
- Le Borgne, J. F., Pello, R., & Sanahuja, B. 1992, A&AS, 95, 87
- Lubin, L. M., & Sandage, A. 2001, AJ, 122, 1084
- Maraston, C. 2005, MNRAS, 362, 799
- Mei, S., et al. 2006, ApJ, 644, 759
- Moran, S. M., Ellis, R. S., Treu, T., Smail, I., Dressler, A., Coil, A. L., & Smith, G. P. 2005, ApJ, 634, 977
- Nagamine, K., Cen, R., Hernquist, L., Ostriker, J. P., & Springel, V. 2005, ApJ, 627, 608
- Nagashima, M., Lacey, C. G., Okamoto, T., Baugh, C. M., Frenk, C. S., & Cole, S. 2005, MNRAS, 363, L31
- Pahre, M. A., Djorgovski, S. G., & de Carvalho, R. R. 1996, ApJ, 456, L79
- Papovich, C., et al. 2006, ApJ, 640, 92
- Quadri, R., et al. 2007, ApJ, 654, 138
- Reddy, N. A., Erb, D. K., Steidel, C. C., Shapley, A. E., Adelberger, K. L., & Pettini, M. 2005, ApJ, 633, 748
- Robertson, B., Cox, T. J., Hernquist, L., Franx, M., Hopkins, P. F., Martini, P., & Springel, V. 2006, ApJ, 641, 21
- Rubin, K. H. R., van Dokkum, P. G., Coppi, P., Johnson, O., F rster Schreiber, N. M., Franx, M., & van der Werf, P. 2004, ApJ, 613, L5
- Rusin, D., & Kochanek, C. S. 2005, ApJ, 623, 666
- Rusin, D., et al. 2003, ApJ, 587, 143
- Salpeter, E. E. 1955, ApJ, 121, 161
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
- Schneider, R., Ferrara, A., Natarajan, P., & Omukai, K. 2002, ApJ, 571, 30
- Sirianni, M. J. J., et al. 2005, PASP, 117, 1049
- Smail, I., Dressler, A., Couch, W. J., Ellis, R. S., Oemler, A. J., Butcher, H., & Sharples, R. M. 1997, ApJS, 110, 213
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., Erb, D. K., Reddy, N. A., & Pettini, M. 2005, ApJ, 626, 44
- Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
- Tinsley, B. M. 1980, *Fundam. Cosm. Phys.*, 5, 287
- Tonry, J., & Davis, M. 1979, AJ, 84, 1511
- Tran, K.-V. H., van Dokkum, P., Franx, M., Illingworth, G. D., Kelson, D. D., & Schreiber, N. M. F. 2005a, ApJ, 627, L25
- Tran, K.-V. H., van Dokkum, P., Illingworth, G. D., Kelson, D., Gonzalez, A., & Franx, M. 2005b, ApJ, 619, 134
- Treu, T., Ellis, R. S., Liao, T. X., & van Dokkum, P. G. 2005a, ApJ, 622, L5
- Treu, T., Stiavelli, M., Casertano, S., Moller, P., & Bertin, G. 1999, MNRAS, 308, 1037

- Treu, T., Stiavelli, M., Casertano, S., Moller, P., & Bertin, G. 2002, *ApJ*, 564, L13
- Treu, T., et al. 2005b, *ApJ*, 633, 174
- van de Ven, G., van Dokkum, P. G., & Franx, M. 2003, *MNRAS*, 344, 924
- van der Marel, R. P., & van Dokkum, P. G. 2006a, *ApJ*, submitted (astro-ph/0611571)(vdMvD06a)
- . 2006b, *ApJ*, submitted (astro-ph/0611577)(vdMvD06b)
- van der Wel, A., Franx, M., van Dokkum, P. G., & Rix, H.-W. 2004, *ApJ*, 601, L5
- van der Wel, A., Franx, M., van Dokkum, P. G., Rix, H.-W., Illingworth, G. D., & Rosati, P. 2005, *ApJ*, 631, 145
- van Dokkum, P. G. 2001, *PASP*, 113, 1420
- . 2005, *AJ*, 130, 2647
- van Dokkum, P. G., & Ellis, R. S. 2003, *ApJ*, 592, L53
- van Dokkum, P. G., & Franx, M. 1996, *MNRAS*, 281, 985
- . 2001, *ApJ*, 553, 90
- van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 1998a, *ApJ*, 504, L17
- van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 2001, *ApJ*, 553, L39
- van Dokkum, P. G., Franx, M., Kelson, D. D., Illingworth, G. D., Fisher, D., & Fabricant, D. 1998b, *ApJ*, 500, 714
- van Dokkum, P. G., & Stanford, S. A. 2003, *ApJ*, 585, 78
- van Dokkum, P. G., et al. 2003, *ApJ*, 587, L83
- . 2006, *ApJ*, 638, L59
- Webb, T. M. A., et al. 2006, *ApJ*, 636, L17
- Worthey, G. 1994, *ApJS*, 95, 107
- Worthey, G., Faber, S. M., & Gonzalez, J. J. 1992, *ApJ*, 398, 69
- Wuyts, S., van Dokkum, P. G., Kelson, D. D., Franx, M., & Illingworth, G. D. 2004, *ApJ*, 605, 677
- Yan, H., et al. 2004, *ApJ*, 616, 63
- Ziegler, B. L., Bower, R. G., Smail, I., Davies, R. L., & Lee, D. 2001, *MNRAS*, 325, 1571